

STIC-Biotech/ChemLib

180560

mg

From: Swope, Sheridan
Sent: Saturday, February 25, 2006 12:04 AM
To: STIC-Biotech/ChemLib
Subject: 09/917,376

For 09/917,376, pls intereference search:

SID 1 against the NT and AA databases

SID 3 against the NT and AA databases

SID 4 against the NT and AA databases

SID 5 against the NT and AA databases

RECEIVED
FEB 27 2006
STIC

Thanks!

Sheridan Swope, Ph.D.
Patent Examiner, AU 1656
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E03C70 Remsen Bld (Mailbox)

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 180560

TO: Sheridan Swope
Location: rem/2B71/3C70
Art Unit: 1656
Friday, March 03, 2006
Case Serial Number: 09/917376

From: Kristine Hensle
Location: Biotech-Chem Library
REM-1B69
Phone: (571) 272-4161

Kristine. Hensle@uspto.gov

Search Notes

Examiner Swope,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Kristine Hensle
Librarian
STIC Biotech/Chem Library
(571)272-4161

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 13:55:49 ; Search time 55.1526 Seconds
(without alignments)
1434.576 Million cell updates/sec

Title: US-09-917-376-1

Perfect score: 5134

Sequence: 1 MDRSENRLTWRRRLVLL.....RASFGSVNPATPTADTYLQX 957

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5-COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6-COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H-COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PGTUS-COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE-COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	334	6.5	1751	2	US-09-136-574A-44	Sequence 44, Appl
2	301.5	5.9	1749	2	US-09-640-419C-28	Sequence 28, Appl
3	288	5.6	96	2	US-09-119-507B-112	Sequence 112, Appl
4	288	5.6	96	2	US-09-547-693-236	Sequence 236, Appl
5	276	5.4	1426	2	US-09-136-574A-43	Sequence 43, Appl
6	260	5.1	2736	2	US-09-252-981A-30227	Sequence 30227, A
7	234	4.6	2137	2	US-09-134-001C-4463	Sequence 4463, Ap
8	225.5	4.4	551	1	US-09-033-537A-1	Sequence 1, Appl
9	221.5	4.3	258	2	US-10-104-047-3034	Sequence 3034, Ap
10	221	4.3	2283	2	US-10-172-502-4	Sequence 4, Appl
11	219	4.3	1236	2	US-09-769-787-109	Sequence 109, Appl
12	216.5	4.2	8991	2	US-08-714-741-32	Sequence 32, Appl
13	215.5	4.2	76	2	US-09-547-693-233	Sequence 233, Appl
14	213	4.1	3892	2	US-09-328-352-5503	Sequence 5503, Ap
15	210.5	4.1	521	1	US-08-276-213-3	Sequence 3, Appl
16	209	4.1	2315	2	US-09-949-002-314	Sequence 314, Appl
17	209	4.1	2415	2	US-09-949-002-398	Sequence 398, Appl
18	208.5	4.1	206	2	US-08-529-055-54	Sequence 54, Appl
19	208.5	4.1	493	2	US-09-198-956-10	Sequence 10, Appl
20	208.5	4.1	493	2	US-09-198-955A-12	Sequence 12, Appl
21	208.5	4.1	493	2	US-09-694-531-12	Sequence 12, Appl
22	208.5	4.1	493	2	US-09-670-141-10	Sequence 10, Appl
23	208.5	4.1	493	2	US-10-072-152-12	Sequence 12, Appl
24	205	4.0	918	2	US-09-200-650E-1	Sequence 1, Appl
25	204.5	4.0	1970	2	US-09-538-092-1005	Sequence 1005, Ap
26	202.5	3.9	476	2	US-09-339-159B-4	Sequence 4, Appl
27	200.5	3.9	700	1	US-07-862-588B-2	Sequence 2, Appl

28	200	3.9	183	2	US-08-529-055-50	Sequence 50, Appl
29	199.5	3.9	490	2	US-09-109-841-2	Sequence 2, Appl
30	199.5	3.9	616	2	US-09-136-574A-47	Sequence 47, Appl
31	199.5	3.9	1367	2	US-09-487-558B-108	Sequence 108, Appl
32	197.5	3.8	1209	2	US-09-949-002-493	Sequence 493, Appl
33	197.5	3.8	1209	2	US-09-949-002-494	Sequence 494, Appl
34	197	3.8	1060	2	US-08-911-393-2	Sequence 2, Appl
35	197	3.8	1060	2	US-09-955-909-2	Sequence 2, Appl
36	192.5	3.7	271	2	US-09-248-796A-21050	Sequence 21050, A
37	191.5	3.7	412	1	US-08-313-288B-18	Sequence 18, Appl
38	191	3.7	105	2	US-09-547-693-230	Sequence 230, Appl
39	191	3.7	2169	2	US-09-949-016-6930	Sequence 6930, Ap
40	190	3.7	1092	2	US-09-147-408B-15	Sequence 15, Appl
41	189.5	3.7	1187	2	US-09-949-016-6513	Sequence 6513, Ap
42	189	3.7	933	2	US-08-293-728-2	Sequence 2, Appl
43	189	3.7	933	2	US-09-421-868-2	Sequence 2, Appl
44	188.5	3.7	193	2	US-08-529-055-49	Sequence 49, Appl
45	187	3.6	936	2	US-08-956-171E-5249	Sequence 5249, Ap

ALIGNMENTS

RESULT 1

US-09-136-574A-44
; Sequence 44, Application US/09136574A
; Patent No. 6294366
; GENERAL INFORMATION:
; APPLICANT: Farrington, Graham K.
; Anderson, Paige
; Gibbs, Moreland
; Bergquist, Peter
; Daniels, Roy
; Morgan, Hugh W.
; Williams, Diane P.

TITLE OF INVENTION: Compositions and Methods for Treating Cellulose Containing Cellulase Enzyme Compositions

NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Center, P.O. Box 457
CITY: Spring House
STATE: PA
COUNTRY: USA
ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/136,574A

FILING DATE: 19-Aug-1998

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/932,571

FILING DATE: September 19, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Bak, Mary E.

REGISTRATION NUMBER: 31,215

REFERENCE/DOCKET NUMBER: 1997US001/CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-540-9200

TELEFAX: 215-540-5818

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 44:

SEQUENCE CHARACTERISTICS:

LENGTH: 1751 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 44:

Fabrics Using Truncated

US-09-136-574A-44

Query Match 6.5%; Score 334; DB 2; Length 1751;
Best Local Similarity 21.2%; Pred. No. 2.5e-13;
Matches 219; Conservative 125; Mismatches 314; Indels 376; Gaps 56;

Qy 18 SLAATASPAVAALGVLPATITASPAAH-----TTQP---YTWNSVAI 59
Db 186 SVVAET-----AASLAVASVVIKERNQSKAASYLQAHAKDLFEFADTTRSDAGTYATGFY 240
Qy 60 GCGGFVDGIVFNEGAPGILYRTD-----ICGMVW-----DAANGRWIP 99
Db 241 TSGGFIDDLG---AAWLYIATNDSSYLTKABELMSEYANGTWTQCDWDDVRYGLIM 297
Qy 100 LLDVWGNNGVYVSIADPINTNKVAAVGM-YTNSWDPNMGAILRSSDQ-----152
Db 298 LAKITGKEL--YKGAVERNLDHWTDRITYTPKGMAYLTGW---GSLRYATTAAFLACVY 351
Qy 153 ATWQ-----ITPLPKLGNMPGR-----GMGERLAVDPNNNIIYFGAPSG 194
Db 352 ADWSGCDSNKTKYLNFASKQIDYALGST--GRSFVVGFGTNPQPHPHRN-----400
Qy 195 KGLWRSTDGATWSQMTNEPD-----VGTYIANPTDTTGVQSDIQGVW--VAFDKSSS 246
Db 401 -----AHSSWANSKIPDYERHILYLGALVGGPGSDSYNDIDITYVONEVACDYNAG 452
Qy 247 SLGQASKTIFVGVADPNNPVFWSRDRGATQWQAVPGAPTGFPHKGVDPVNVHLYIAT--304
Db 453 IVGALAK-----MYQLYGGEPID--DFKAIETPTNDEIFVESKF 489
Qy 305 SNTGGPYDSSGVDWKFVSVTSGTWTTRISVPVSTDTAN-DYF-----GYSGLTIDRQ 354
Db 490 GNSQGP--NYTEVSIYVNRITGW---PPRVTDKLSFKYFIDITELIQAGYS-----535
Qy 355 HPNTIMVATOISWPPDTIIFRSDGGATWTRIDWTSYPNRSRYVL-DISASEPWLTFGV 413
Db 536 -PDVVKVDT-----YYIEGKISGYPVWD--KRNIIYVLVDPSGTK-----I 575
Qy 414 QNPVPVSPKLGW---MDEAMAIIDPNSDRMLYGTGATLYATNDLTWKDSCGQIHIAPM 469
Db 576 YPGGEVHKQAOFKLSVPGYFWDPTN-DPSYKGLTSQLEKKNYIAAYDNN--L 628
Qy 470 VKGLETAVNDLSP-----PSGAPLISALGDLGGFTHADYVAVPSTIFT---SPVFT 519
Db 629 VMGLEPCAATSTAPTSTPTPTPTPTVTA-----TPTPTPTPTGSP--G 674
Qy 520 TGTSDVYAEALNPISIIVRAGSFDSSQPNDRHAFSTDCGNWFGQSGPEGVTTCGTVAAS 579
Db 675 TSGGVKLVLYKNNETSASTGSRP-----WPK-----IVNGG--SSS 708
Qy 580 ADGSRF---VW--APGDPGPVYVAVGFGNSWAASQGVPAQAIRSDRVNPKTFYALSN 634
Db 709 VDLRVRKIRYWTVDGDKQSAVC-----DW-----AOLGASNVT-FNFVKLSG 752
Qy 635 TFYRSTDDGVTFOVAAAGLPSSGAVGYMHAFVPEKE-GDLWLAASSGLYHSTNGGSSWA 693
Db 753 V-----SGADYY-LEVGP-SSGAGLQ-----FGKOTGDIQVRFKNKDSNYNQADWSW 800
Qy 694 ITGVSAVNVGFKSAPGSSYPVAVFVGTGTGAYRSDDCGTTWVLINDDOHQYQNWG 753
Db 801 LOSMTN-----YGENAKVTLY-----VDG-----VLV-----WG 824
Qy 754 QAITGDHANLRVYICTNGRGIVYDGGAPSGSPSPSVSPSASPSPSPSPSPSPSPS 813
Db 825 Q-----BFGGA-----TPAPTSTATPT 841
Qy 814 PSPSSSP 873
Db 842 PPTAPT 887
Qy 874 KVQYKNDSPAGDNOIKPGLQVNTGSSSYDLSTVTYVYFTDRGSSSTLVYNCDMAAIG 933
Db 888 KVMYANGNLSSPTNVLNPKIENVGTTAVDLRVRKYVWYTTIDGEATQSV-----938

Qy 934 CGNIRASFGSVNPA 947
Db 939 -----SVASSINPA 947

RESULT 2
US-09-640-419C-28
Sequence 28, Application US/09640419C
Patent No. 6630615
GENERAL INFORMATION:
APPLICANT: Bidney, Dennis L
APPLICANT: Crasta, Oswald R
APPLICANT: Hu, Xu
APPLICANT: Lu, Guihua
TITLE OF INVENTION: DEFENSE-RELATED SIGNALING GENES AND METHODS OF USE
CURRENT APPLICATION NUMBER: US/09/640,419C
CURRENT FILING DATE: 2000-08-17
PRIOR APPLICATION NUMBER: 60/149,656
PRIOR FILING DATE: 1999-08-18
PRIOR APPLICATION NUMBER: 60/206,405
PRIOR FILING DATE: 2000-05-23
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 28
LENGTH: 1749
TYPE: PRT
ORGANISM: Synechocystis PCC6803
US-09-640-419C-28

Query Match 5.9%; Score 301.5; DB 2; Length 1749;
Best Local Similarity 19.0%; Pred. No. 3.5e-11;
Matches 225; Conservative 117; Mismatches 402; Indels 439; Gaps 46;

Qy 17 VSLAATASPAVAALGVLPATITASPAAHATTPYTWSNVAIGGG-----62
Db 23 LALLACSSPSG---NVLAQNTAPADGTGTTVDACGNQFNIGGSLSGDQNLPHSLQ 78
Qy 63 -----GFVDGIVNEGAPGILYRTDIDGMYRWA 92
Db 79 QFGLDQGIANFLSNPDIRNLTIRIVGDASIIINGLIQVSGGNANFLMNPAGMIFGPN 138
Qy 93 A-----NGRWIPLLDVGNWNGVGVVS-----116
Db 139 SINVPDVFVTTGSAIGFNDQFV-----PSNDYNALIGNPSQAFADLANPGLIINA 193
Qy 117 -----IAADPINTNKYMAAVGMVYTNWDNDGAILRSS-----149
Db 194 GDLSTEGKNLTFLAGNIVNTGSLAAPGNGITVAAPGQNRIRISQAGSLLSLEVEVSQ 253
Qy 150 -DQATWQITPLPKLGNMPGRGNGERLAVDPND-----NIIYFAPSGKGLWRSTD 202
Db 254 MNQGSFVLDPITLLTQGNLNLG--LAVQPNGSVTTNGTNAVSLPLGSVTTISGND 311
Qy 203 SGATWSQMTNFPDVGVTIANPTDTTGYCS---DIQ-----GVWVAFD-KSSSSILGQASK 253
Db 312 AS---GKSTISS--GGQVIAIGDQIAVQGATVDVSGNGGGGTVRIGGDFQGLTLPNASQ 367
Qy 254 TIFVGVAADPNPV-----FWSRDGGA--TWQAVPGAPTGFPHKGVDFPNVHLYIATNT 307
Db 368 TLI-----DSNSVVKADALLTNGGTIVWADDSTFSGNI-----SAQ 406
Qy 308 GGPVDSGGDWKPSVTSCTWTRISPVSTDTANDYFGYGLTIDRQHPNTIMVATQISW 367
Db 407 GGTMGNGG---FVETSG-----AKSLMVDTDARVNTFATMGEIGT 444
Qy 368 W---PDTIIFRSTDDGATWTRWD-----WTSYPRNSRLRYVLDISNEPWLTFGVQNPVP 420
Db 445 WLLDPLEIIVGTTDOLLADPKLSVLVATTITSLDNGNVILQADQSIQAVQANFADPSAP--502
Qy 421 SPKLGWMDAMAIIDPNSDRMLYGTGATLYA-----TNDLTWKDSCGQIHI--466

[illegible]

RESULT 8
US-09-033-537A-1
; Sequence 1, Application US/09033537A
; Patent No. 5958083
; GENERAL INFORMATION:
; APPLICANT: Onishi, Masahiro
; APPLICANT: Fich, Merete
; APPLICANT: Toft, Annette Hanne
; APPLICANT: Sh lein, Martin
; TITLE OF INVENTION: Prevention Of Back-Staining
; TITLE OF INVENTION: In Stone Washing
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5958083o No. 5958083disk of No. 5958083th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: U.S.A.
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/033,537A
; FILING DATE: 02-MAR-1998
; CLASSIFICATION: 008
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 0993/95
; FILING DATE: 08-SEP-1995
; APPLICATION NUMBER: PCT/DK96/00364
; FILING DATE: 03-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4492.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 551 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-033-537A-1

Query Match 4.4%; Score 225.5; DB 1; Length 551;
Best Local Similarity 23.8%; Pred. No. 7.9e-07;
Matches 114; Conservative 72; Mismatches 166; Indels 127;
Gaps 22;

Query Match 4.4%; Score 225.5; DB 1; Length 551;
Best Local Similarity 23.8%; Pred. No. 7.9e-07;
Matches 114; Conservative 72; Mismatches 166; Indels 127; Gaps 22;

[illegible]

RESULT 9

```

US-10-104-047-3034
; Sequence 3034, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3034
; LENGTH: 258
; TYPE: PRt
; ORGANISM: Homo sapiens
US-10-104-047-3034

```

[illegible]

RESULT 10
US-10-172-502-4

```

; Sequence 4, Application US/10172502
; Patent No. 6841154
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy et al.
; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES...
; FILE REFERENCE: P07263JUS01/BAS
; CURRENT APPLICATION NUMBER: US/10/172,502
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/298,098
; PRIOR FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2283
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-10-172-502-4

```

Query Match	4.38;	Score 221;	DB 2;	Length 2283;
Best Local Similarity	20.79;	Pred. No. le-05;		
Matches 178;	Conservative 120;	Mismatches 321;	Indels 240;	Gaps 30;
Qy	145	ILRSSDQG---ATWQITPLPKLGGNMPGRGMGERLAVDPNNNNILYFGAPSGKGLWRST	201	
Db	608	VLITTDNGTGVNTVVTGLP-----SGLSYDSATNSII--GTPTKIG--QST	650	
Qy	202	DSGATWSQMTNFPDVGTYIANPTDITTCYQSDIQGVWVAFDKSSSLGQASKTIFVGVAD	261	
Db	651	VTVVSTDQANN-KSTTTFTTINVVDTTA-----PTVTPIGDQSSEVY-----S	691	
Qy	262	PNNPV---FWSRGGATQWAPGAPTGFPHKGVPDPNVHL-----YIATSNITGGPYDG	313	
Db	692	PIGPIKATQDNGSNVNTVTGLPSGL-----TFDSTNNITSGTPTNIGTSTISIVSTD	746	

[illegible]

```

QY      888  QIRPGL-----QVNTGSSVDLSVTVRVWTFTRDGGSSLTIVNCDWAAI  932
Db      1217  QSAFAFLSELSESTSESVSSSTSESTSLSDST-----SESGSTSTSLGN----- 1265
QY      933  GCGNIRASFGSVNPATPTA 951
Db      1266  -----STSGSTSIISTSTS 1278

RESULT 11
US-09-769-787-109
; Sequence 109, Application US/09769787
; Patent No. 6936252
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 109
; LENGTH: 1236
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-109

```

Qy		716	AVFWGTIGVTCATRSDDCGTTWLINDDQHGYGNWGQAITGDHANLRVVYIGNRGI	775
Dd		488	SASTSASASASTSAGESASTSAS-----ASASTSAGAS---ASTSASGSAS	530
Qy		776	VYGDIGCAPSGSPSVGPSASPSPSSPSPSPSPSPSSPSPPSPSPSPSPSPSP	835
Dd		531	TSTTSASASTSASASASTSASASASTSASESASTTSTASASTSASESATSTAS	590
Qy		836	RSPSPASGSPSSPSPSPSPSSPSSPSPSPSPSPSPSPSPSPSPSPSPSPSP	882
Dd		591	ASASTSASASASTSASASASTSASASTSASESATSTASASTSASASTSASASTS	650
Qy		883	APGDNQIKPLGVLTNGSSSDVLSTVTWRYFTRGSGSTLVNCDWAIGCNTRASFG	942
Dd		651	ASASASTSASVSASTSASASTSASASTSASASTSASESATSTASASTSASASTSASAS	710
Qy		943	SVPNPATPTADT	953
Dd		711	ASTSASESAT	721

```

RESULT 12
US-08-714-741-32
; Sequence 32, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF.
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8991 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-714-741-32
;
Query Match          4.2%; Score 216.5; DB 2; Length 8991;
Best Local Similarity 30.6%; Pred. No. 0.00012;

```

```

Query Match      4.2%; Score 215.5; DB 2; Length 76;
Best Local Similarity 51.8%; Pred. No. 2.7e-07;
Matches 44; Conservative 2; Mismatches 28; Indels 11; Gaps 1;

QY 781 GGAPSGSPSVSPASPSLSPSPSSSPSPSPSPSSSPSSSPSPSPSPSPSPSPSP 840
Db 1 GSSTRASPPP-----PSPPSPPPPPPPPPPPSPPPSPPPSPPPSPPPSPPPSP 49
QY 841 SASPSSSPSSSPSSSPSSSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 865
Db 50 PPSPPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPS 74

RESULT 14
US-09-328-352-5503
; Sequence 5503, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: GARY L. BRETON ET AL.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328.352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5503
; LENGTH: 3892
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5503

Query Match      4.1%; Score 213; DB 2; Length 3892;
Best Local Similarity 20.7%; Pred. No. 6.9e-05;
Matches 238; Conservative 13; Mismatches 425; Indels 348; Gaps 58;

QY 33 GVLPITAITSPAHAAT-TQPYTWSNVAIGGGGFVDG--IVFNEG---APGILYVRTDIGG 86

```

[illegible]

Query Match	4.1%	Score	210.5	DB 1	Length	521																																											
Best Local Similarity	22.2%	Pred.	No. 7.2e-06																																														
Matches	137	Conservative	74	Mismatches	198	Indels	207	Gaps	31																																								
QY	379	GGATWTRIMDWT	SVPNRSRLV	LDISAE	PMLTFG	VQVQNP	PPSPKLG	MM-----	427																																								
Db	2	GGGYW-----	HTSGREILD	AN-----	NVP	RIAGI	NWFG	FETCNYV	VHG 40																																								
QY	428	-----DE	MAIDPNS	DRMLYGT	GATL	YATND	LTKWD	SGGQ	IHIAPM --- VKGLEETA	VND 480																																							
Db	41	LWSR	DRYRSLMD	QIKS--	LGYN	TIRLP	YSDDLK	PGT	MPNSINF	YQMNDLOGL	TSLQVMD 98																																						
QY	481	LISPP	SGAPLIS	ALGD	LG-----	GF	THADV	AVP	STFTIS	PVFTT	GTSDV	VAEL-----	529																																				
Db	99	KI-----	VA	YAGQ	IGLR	IL	DRH	PHD	CSG	QSAL	WYTS	SVSEAT	WISD	LQAL	AQRYK 150																																		
QY	530	NPS	LI	IV	AGS	FDP	SSQ	PND	RH	VAF	ST	DG	GKN	WFG	SGSE	PPG-----	VTTG	TGV	AA 578																														
Db	151	NPT	V-----	G	F	D	LH	N	E	P	H	D	P	A	C	M	G	G	D	P	S	I	D	W	R	L	A	A	E	R	A	G	N	A	V	L	S	N	P	N	L	L	I	F	E	B	G-----	V	Q 204
QY	579	SAD	GS	F	V	A	P	A	G	D	P	G	P	V	V	Y	A	V	G	N	S	W	A	A	S	Q-----	G	V	P-----	A	N	A	O	I	R	S	D	R	N	P	K	T	F	Y 629					

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 14:20:38 ; Search time 184.353 Seconds
(without alignments)
2169.009 Million cell updates/sec

Title: US-09-917-376-1
Perfect score: 5134
Sequence: 1 MDRSENRLTWSRRLVSL.....RASFGSVNPATPTADTYLQX 957

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pcp:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5134	100.0	957	3	US-09-917-376-1
2	5134	100.0	957	4	US-10-155-400-1
3	4036	78.6	740	3	US-09-917-376-3
4	4036	78.6	740	3	US-09-917-376-6
5	4036	78.6	740	4	US-10-155-400-3
6	4036	78.6	740	4	US-10-155-400-6
7	2478	48.3	882	4	US-10-156-761-9395
8	1680	32.7	726	3	US-09-917-376-7
9	1680	32.7	726	4	US-10-155-400-7
10	1625.5	31.7	838	4	US-10-420-191-2
11	1604	31.2	818	4	US-10-026-994-2
12	1442	28.1	739	4	US-10-156-761-10111
13	1125	21.9	812	4	US-10-395-241-12
14	1117	21.8	789	4	US-10-395-241-14
15	1114	21.7	826	4	US-10-395-241-18
16	848	16.5	555	3	US-09-927-827-47
17	823	16.0	1228	3	US-09-917-384-1
18	823	16.0	1228	3	US-09-917-383-1
19	702	13.7	762	3	US-09-917-378-1
20	511.5	10.0	1043	3	US-09-917-384-6
21	511.5	10.0	1043	3	US-09-917-383-6
22	462	9.0	88	3	US-09-917-376-5
23	462	9.0	88	4	US-10-155-400-5
24	462	9.0	89	3	US-09-917-376-4
25	462	9.0	89	4	US-10-155-400-4
26	462	9.0	154	3	US-09-917-378-4
27	459	8.9	150	3	US-09-917-384-5

28	459	8.9	150	3	US-09-917-383-5	Sequence 5, Appli
29	389	7.6	284	5	US-10-418-032-270	Sequence 270, App
30	346.5	6.7	901	5	US-10-939-262-12	Sequence 12, Appl
31	336.5	6.6	741	4	US-10-156-761-8100	Sequence 8100, Ap
32	326.5	6.4	628	5	US-10-939-262-14	Sequence 14, Appl
33	288	5.6	96	4	US-10-437-708-236	Sequence 236, App
34	288	5.6	96	4	US-10-395-402-112	Sequence 112, App
35	288	5.6	96	5	US-10-257-193-236	Sequence 236, App
36	288	5.6	96	5	US-10-418-032-236	Sequence 236, App
37	274	5.3	406	5	US-10-450-763-57609	Sequence 57609, A
38	268	5.2	284	5	US-10-418-032-271	Sequence 271, App
39	262	5.1	2468	4	US-10-246-330-4	Sequence 4, Appli
40	262	5.1	2468	4	US-10-282-122A-66335	Sequence 66335, A
41	261.5	5.1	599	3	US-09-955-555A-29	Sequence 29, Appl
42	253	4.9	2435	4	US-10-282-122A-47453	Sequence 47453, A
43	252.5	4.9	2117	4	US-10-120-801-63	Sequence 63, Appl
44	246	4.8	1066	4	US-10-647-196-43	Sequence 43, Appl
45	243.5	4.7	1049	4	US-10-282-122A-49900	Sequence 49900, A

ALIGNMENTS

RESULT 1

US-09-917-376-1
; Sequence 1, Application US/09917376
; Publication No. US20040038334A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOLYTICUS
; FILE REFERENCE: 40197.4US01
; CURRENT APPLICATION NUMBER: US/09/917,376
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 957
; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (957)
; OTHER INFORMATION: Any amino acid
US-09-917-376-1

Query Match	100.0%	Score 5134;	DB 3;	Length 957;
Best Local Similarity	100.0%	Pred. No. 3.9e-280;	Mismatches 0;	Indels 0;
Matches 956;	Conservative 0;			Gaps 0;
Qy	1	MDRSENRLTWSRRLVSLAATASFAVAALGVLPALITASPAHAATTPYTSNVAIG	60	
Db	1	MDRSENRLTWSRRLVSLAATASFAVAALGVLPALITASPAHAATTPYTSNVAIG	60	
Qy	61	GGGFVDGIVFNEGAPGILYVRTDIGMYRWDAAANGRWIPLLDWGNNWNGYGVVSIAD	120	
Db	61	GGGFVDGIVFNEGAPGILYVRTDIGMYRWDAAANGRWIPLLDWGNNWNGYGVVSIAD	120	
Qy	121	PINTNKWAAVGMVYNSWDNDGAILRSSDQCATWQITLPEKLGNNMGRGNGERLAVD	180	
Db	121	PINTNKWAAVGMVYNSWDNDGAILRSSDQCATWQITLPEKLGNNMGRGNGERLAVD	180	
Qy	181	PNNNLIYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVA	240	
Db	181	PNNNLIYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVA	240	
Qy	241	FDKSSSSILQASKTIFVGVADPNPNFVMSRDGGATWQAVPGAPTGFIPHKGVDPVNVHL	300	
Db	241	FDKSSSSILQASKTIFVGVADPNPNFVMSRDGGATWQAVPGAPTGFIPHKGVDPVNVHL	300	

; APPLICANT: ADNEY, WILLIAM S.
 ; APPLICANT: VINZANT, TODD B.
 ; APPLICANT: HIMMEL, MICHAEL E.
 ; TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
 ; TITLE OF INVENTION: CELLULOYTICUS
 ; FILE REFERENCE: 40197.4US01
 ; CURRENT APPLICATION NUMBER: US/09/917,376
 ; CURRENT FILING DATE: 2001-07-28
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 740
 ; TYPE: PRT
 ; ORGANISM: Acidothermus cellulolyticus
 ; FEATURE:
 ; OTHER INFORMATION: Catalytic domain GH74
 US-09-917-376-3

Query Match 78.6%; Score 4036; DB 3; Length 740;
 Best Local Similarity 100.0%; Pred. No. 1.5e-218;
 Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 47 ATTQPYTWSNVAIGGGFVDGIVNEGAPGILYVRTDIGMYRWDAAANGRWIPLLDWVGW 106
 Db 1 ATTQPYTWSNVAIGGGFVDGIVNEGAPGILYVRTDIGMYRWDAAANGRWIPLLDWVGW 60
 QY 107 NNWGVGVVSIADPINTNKVAAVGMVYNSWDNDGAILRSSDQATWQITPLPFLKGG 166
 Db 61 NNWGVGVVSIADPINTNKVAAVGMVYNSWDNDGAILRSSDQATWQITPLPFLKGG 120
 QY 167 NMPGRGGERLAVDPNNNDNLYFGAPSGKGLWRSTDGATWSQMTNFPDVGTYIANPTDT 226
 Db 121 NMPGRGGERLAVDPNNNDNLYFGAPSGKGLWRSTDGATWSQMTNFPDVGTYIANPTDT 180
 QY 227 TGYQSDIQGVVWAFDKSSSLGQASKTIIFGVADPNPNVFWSDGGATWQAVPGATGF 286
 Db 181 TGYQSDIQGVVWAFDKSSSLGQASKTIIFGVADPNPNVFWSDGGATWQAVPGATGF 240
 QY 287 IPHKGVPDPVNHVLIATNTGGPYDGGSDGVKFSVTSGTWTRISPVPSDTTANDYFGY 346
 Db 241 IPHKGVPDPVNHVLIATNTGGPYDGGSDGVKFSVTSGTWTRISPVPSDTTANDYFGY 300
 QY 347 SGLTIDRQHPNTIMVATQISWMPDTIIFRSTDGATWTRIMDWTSPNRSRLRYVLDISAE 406
 Db 301 SGLTIDRQHPNTIMVATQISWMPDTIIFRSTDGATWTRIMDWTSPNRSRLRYVLDISAE 360
 QY 407 PWTFTGVQPNPPVSPKLGWDEMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI 466
 Db 361 PWTFTGVQPNPPVSPKLGWDEMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI 420
 QY 467 APMVKGLEETAANDLISPPSGAPLISALGDLGGFTHADVTAVPSTIETSPVFTTGTSDY 526
 Db 421 APMVKGLEETAANDLISPPSGAPLISALGDLGGFTHADVTAVPSTIETSPVFTTGTSDY 480
 QY 527 AELNPSIIVRAGSDPPSQPNDRHVAFTDGGKWNFGQSEPGVTTGCTVAASADGSRFV 586
 Db 481 AELNPSIIVRAGSDPPSQPNDRHVAFTDGGKWNFGQSEPGVTTGCTVAASADGSRFV 540
 QY 587 WAPGDPQPVVYAVFGNSWAASQGVPANAIQIRSDRVNPKTFYALSNGTYRSTDDGVTF 646
 Db 541 WAPGDPQPVVYAVFGNSWAASQGVPANAIQIRSDRVNPKTFYALSNGTYRSTDDGVTF 600
 QY 647 QPVAAGLPSSGAVGVNMFHAPGKEDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG 706
 Db 601 QPVAAGLPSSGAVGVNMFHAPGKEDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG 660
 QY 707 KSAFGSSYPVAVVGTIGGVTGAYRSDCCGTWVLINDDOHQYGNWGOAITGDHANLRRV 766
 Db 661 KSAFGSSYPVAVVGTIGGVTGAYRSDCCGTWVLINDDOHQYGNWGOAITGDHANLRRV 720
 QY 767 YITNGRGIVYDGGAPSG 786
 Db 721 YITNGRGIVYDGGAPSG 740

RESULT 4
 US-09-917-376-6
 ; Sequence 6, Application US/09917376
 ; Publication No. US20040038334A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DING, SHI-YOU
 ; APPLICANT: ADNEY, WILLIAM S.
 ; APPLICANT: VINZANT, TODD B.
 ; APPLICANT: HIMMEL, MICHAEL E.
 ; TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
 ; TITLE OF INVENTION: CELLULOYTICUS
 ; FILE REFERENCE: 40197.4US01
 ; CURRENT APPLICATION NUMBER: US/09/917,376
 ; CURRENT FILING DATE: 2001-07-28
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 740
 ; TYPE: PRT
 ; ORGANISM: Acidothermus cellulolyticus
 US-09-917-376-6

Query Match 78.6%; Score 4036; DB 3; Length 740;
 Best Local Similarity 100.0%; Pred. No. 1.5e-218;
 Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 47 ATTQPYTWSNVAIGGGFVDGIVNEGAPGILYVRTDIGMYRWDAAANGRWIPLLDWVGW 106
 Db 1 ATTQPYTWSNVAIGGGFVDGIVNEGAPGILYVRTDIGMYRWDAAANGRWIPLLDWVGW 60
 QY 107 NNWGVGVVSIADPINTNKVAAVGMVYNSWDNDGAILRSSDQATWQITPLPFLKGG 166
 Db 61 NNWGVGVVSIADPINTNKVAAVGMVYNSWDNDGAILRSSDQATWQITPLPFLKGG 120
 QY 167 NMPGRGGERLAVDPNNNDNLYFGAPSGKGLWRSTDGATWSQMTNFPDVGTYIANPTDT 226
 Db 121 NMPGRGGERLAVDPNNNDNLYFGAPSGKGLWRSTDGATWSQMTNFPDVGTYIANPTDT 180
 QY 227 TGYQSDIQGVVWAFDKSSSLGQASKTIIFGVADPNPNVFWSDGGATWQAVPGATGF 286
 Db 181 TGYQSDIQGVVWAFDKSSSLGQASKTIIFGVADPNPNVFWSDGGATWQAVPGATGF 240
 QY 287 IPHKGVPDPVNHVLIATNTGGPYDGGSDGVKFSVTSGTWTRISPVPSDTTANDYFGY 346
 Db 241 IPHKGVPDPVNHVLIATNTGGPYDGGSDGVKFSVTSGTWTRISPVPSDTTANDYFGY 300
 QY 347 SGLTIDRQHPNTIMVATQISWMPDTIIFRSTDGATWTRIMDWTSPNRSRLRYVLDISAE 406
 Db 301 SGLTIDRQHPNTIMVATQISWMPDTIIFRSTDGATWTRIMDWTSPNRSRLRYVLDISAE 360
 QY 407 PWTFTGVQPNPPVSPKLGWDEMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI 466
 Db 361 PWTFTGVQPNPPVSPKLGWDEMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI 420
 QY 467 APMVKGLEETAANDLISPPSGAPLISALGDLGGFTHADVTAVPSTIETSPVFTTGTSDY 526
 Db 421 APMVKGLEETAANDLISPPSGAPLISALGDLGGFTHADVTAVPSTIETSPVFTTGTSDY 480
 QY 527 AELNPSIIVRAGSDPPSQPNDRHVAFTDGGKWNFGQSEPGVTTGCTVAASADGSRFV 586
 Db 481 AELNPSIIVRAGSDPPSQPNDRHVAFTDGGKWNFGQSEPGVTTGCTVAASADGSRFV 540
 QY 587 WAPGDPQPVVYAVFGNSWAASQGVPANAIQIRSDRVNPKTFYALSNGTYRSTDDGVTF 646
 Db 541 WAPGDPQPVVYAVFGNSWAASQGVPANAIQIRSDRVNPKTFYALSNGTYRSTDDGVTF 600
 QY 647 QPVAAGLPSSGAVGVNMFHAPGKEDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG 706
 Db 601 QPVAAGLPSSGAVGVNMFHAPGKEDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG 660
 QY 707 KSAFGSSYPVAVVGTIGGVTGAYRSDCCGTWVLINDDOHQYGNWGOAITGDHANLRRV 766

Db 661 KSAFGSSYPVAVVGTGGTGAIRSDCGTTWVLINDDQHOYGNWQAITGDHANLRV 720
Qy 767 YIGTNGRGIVYDGGAPSG 786
Db 721 YIGTNGRGIVYDGGAPSG 740
RESULT 5
US-10-155-400-3
; Sequence 3, Application US/10155400
; Publication No. US20030108988A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
; FILE REFERENCE: NREL 01-36A
; CURRENT APPLICATION NUMBER: US/10/155,400
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
; FEATURE:
; OTHER INFORMATION: Catalytic domain GH74
US-10-155-400-3
Query Match 78.6%; Score 4036; DB 4; Length 740;
Best Local Similarity 100.0%; Pred. No. 1.5e-218;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 47 ATTQPTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGMYRWDANGRWIPLLDWVGW 106
Db 1 ATTQPTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGMYRWDANGRWIPLLDWVGW 60
Qy 107 NMGYNGVWSIAADPINTNKVAAVGMVYTNWDNDGAILRSSDQATWQITPLPKLGG 166
Db 61 NMGYNGVWSIAADPINTNKVAAVGMVYTNWDNDGAILRSSDQATWQITPLPKLGG 120
Qy 167 NMPGRGGERLAVDPNNNDILYFGAPSGKGLWRSTDGATWTRLDWTSYPNRSLRVLDISAE 226
Db 121 NMPGRGGERLAVDPNNNDILYFGAPSGKGLWRSTDGATWTRLDWTSYPNRSLRVLDISAE 180
Qy 227 TGYQSDIQGVVWVAFDKSSSLGQASKTIIFGVADPNPNVFWSRDGGATWQAVPGAPTGF 286
Db 181 TGYQSDIQGVVWVAFDKSSSLGQASKTIIFGVADPNPNVFWSRDGGATWQAVPGAPTGF 240
Qy 287 IPHKGVPDPVNHVLYIATNTGGPYDGGSDGVKFSVTSCTWTRISPVSTDTANDYFGY 346
Db 241 IPHKGVPDPVNHVLYIATNTGGPYDGGSDGVKFSVTSCTWTRISPVSTDTANDYFGY 300
Qy 347 SGLTIDRQHPNTIMVATQISWNPDTIIFRSTDGATWTRLDWTSYPNRSLRVLDISAE 406
Db 301 SGLTIDRQHPNTIMVATQISWNPDTIIFRSTDGATWTRLDWTSYPNRSLRVLDISAE 360
Qy 407 PWLTFGVQPNPPVPSPKLGMWDEAMADPNSDRMLYGTGATLYATNDLTKWDSGGQIHI 466
Db 361 PWLTFGVQPNPPVPSPKLGMWDEAMADPNSDRMLYGTGATLYATNDLTKWDSGGQIHI 420
Qy 467 APWKGLEETA VNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVTTCTSVDY 526
Db 421 APWKGLEETA VNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVTTCTSVDY 480
Qy 527 AELNPSIIVRAGSFPDSSQNDRHVAFSTDGKKNWFQSGEPGGVTTGGTVAASADGRFV 586
Db 481 AELNPSIIVRAGSFPDSSQNDRHVAFSTDGKKNWFQSGEPGGVTTGGTVAASADGRFV 540
Qy 587 WAPGDPGPVAVVGFNSWAAASQGVPAQAIRSDRVNPKTFYALNSGTFFYRSTDGGVTF 646

Db 541 WAPGDPGPVAVVGFNSWAAASQGVPAQAIRSDRVNPKTFYALNSGTFFYRSTDGGVTF 600
Qy 647 QPVAAGLPSSGAVGMFHAVPKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG 706
Db 601 QPVAAGLPSSGAVGMFHAVPKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG 660
Qy 707 KSAFGSSYPVAVVGTGGTGAIRSDCGTTWVLINDDQHOYGNWQAITGDHANLRV 766
Db 661 KSAFGSSYPVAVVGTGGTGAIRSDCGTTWVLINDDQHOYGNWQAITGDHANLRV 720
Qy 767 YIGTNGRGIVYDGGAPSG 786
Db 721 YIGTNGRGIVYDGGAPSG 740
RESULT 6
US-10-155-400-6
; Sequence 6, Application US/10155400
; Publication No. US20030108988A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
; FILE REFERENCE: NREL 01-36A
; CURRENT APPLICATION NUMBER: US/10/155,400
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
US-10-155-400-6
Query Match 78.6%; Score 4036; DB 4; Length 740;
Best Local Similarity 100.0%; Pred. No. 1.5e-218;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 47 ATTQPTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGMYRWDANGRWIPLLDWVGW 106
Db 1 ATTQPTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGMYRWDANGRWIPLLDWVGW 60
Qy 107 NMGYNGVWSIAADPINTNKVAAVGMVYTNWDNDGAILRSSDQATWQITPLPKLGG 166
Db 61 NMGYNGVWSIAADPINTNKVAAVGMVYTNWDNDGAILRSSDQATWQITPLPKLGG 120
Qy 167 NMPGRGGERLAVDPNNNDILYFGAPSGKGLWRSTDGATWTRLDWTSYPNRSLRVLDISAE 226
Db 121 NMPGRGGERLAVDPNNNDILYFGAPSGKGLWRSTDGATWTRLDWTSYPNRSLRVLDISAE 180
Qy 227 TGYQSDIQGVVWVAFDKSSSLGQASKTIIFGVADPNPNVFWSRDGGATWQAVPGAPTGF 286
Db 181 TGYQSDIQGVVWVAFDKSSSLGQASKTIIFGVADPNPNVFWSRDGGATWQAVPGAPTGF 240
Qy 287 IPHKGVPDPVNHVLYIATNTGGPYDGGSDGVKFSVTSCTWTRISPVSTDTANDYFGY 346
Db 241 IPHKGVPDPVNHVLYIATNTGGPYDGGSDGVKFSVTSCTWTRISPVSTDTANDYFGY 300
Qy 347 SGLTIDRQHPNTIMVATQISWNPDTIIFRSTDGATWTRLDWTSYPNRSLRVLDISAE 406
Db 301 SGLTIDRQHPNTIMVATQISWNPDTIIFRSTDGATWTRLDWTSYPNRSLRVLDISAE 360
Qy 407 PWLTFGVQPNPPVPSPKLGMWDEAMADPNSDRMLYGTGATLYATNDLTKWDSGGQIHI 466
Db 361 PWLTFGVQPNPPVPSPKLGMWDEAMADPNSDRMLYGTGATLYATNDLTKWDSGGQIHI 420
Qy 467 APWKGLEETA VNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVTTCTSVDY 526
Db 421 APWKGLEETA VNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVTTCTSVDY 480

QY 527 AELNPSIIVRAGSDPSSQPNDRHVAFTDGGKWFQSGEPGCVTTGGTVAASADGSRFV 586
DB 481 AELNPSIIVRAGSDPSSQPNDRHVAFTDGGKWFQSGEPGCVTTGGTVAASADGSRFV 540
QY 587 WAPGDPGPVVYAVGFGNSWAASQGVANAOIRSDRVNPKTFYALNSGTFFYRSTDDGGVTF 646
DB 541 WAPGDPGPVVYAVGFGNSWAASQGVANAOIRSDRVNPKTFYALNSGTFFYRSTDDGGVTF 600
QY 647 QPVAAGLPSSGAVGMFHAVPKGEGDLMLAASSGLYHSTNGSSWSAITGVSSAVNVGFG 706
DB 601 QPVAAGLPSSGAVGMFHAVPKGEGDLMLAASSGLYHSTNGSSWSAITGVSSAVNVGFG 660
QY 707 KSAPGSSYPVAVGTVGTVGAYRSDCGTTWVLINDDQHQYGNWGOAITGDHANLRV 766
DB 661 KSAPGSSYPVAVGTVGTVGAYRSDCGTTWVLINDDQHQYGNWGOAITGDHANLRV 720
QY 767 YIGTNGRGIVYGDIGGAPSG 786
DB 721 YIGTNGRGIVYGDIGGAPSG 740

RESULT 7

US-10-156-761-9395

; Sequence 9395, Application US/10156761

; Publication No. US20030119018A1

; GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI

; APPLICANT: IKEDA, HARUO

; APPLICANT: ISHIKAWA, JUN

; APPLICANT: HORIKAWA, HIROSHI

; APPLICANT: SHIBA, TADAYOSHI

; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-262

; CURRENT APPLICATION NUMBER: US/10/156,761

; CURRENT FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: JP 2001-204089

; PRIOR FILING DATE: 2001-05-30

; PRIOR APPLICATION NUMBER: JP 2001-272697

; PRIOR FILING DATE: 2001-08-02

; NUMBER OF SEQ ID NOS: 15109

; SEQ ID NO 9395

; LENGTH: 882

; TYPE: PRT

; ORGANISM: Streptomyces avermitilis

US-10-156-761-9395

Query Match 48.3%; Score 2478; DB 4; Length 882;
Best Local Similarity 50.4%; Pred. No. 7.1e-131;
Matches 473; Conservative 128; Mismatches 227; Indels 110; Gaps 14;

QY 23 TASFAVAAL--GVLPIAITASPAHAATQPYTWSNVAIGGGFVDGIVFNEGAGIILYV 80
DB 8 TAVLAALAGLPAGTPPALAASAPTATIAATYSKNAARVDGFGFVGVFVFNSEKNLAYA 67
QY 81 RTDIGMTWDAANGRWITPLLDWGNWNGVVSIAADPINTNKVMAAVGMVYNSWDP 140
DB 68 RTDIGAYRWAESSKTWTPLLDSVGSWDGHTGVVSLASDSVDPNKVAAVGVYTNWDP 127
QY 141 NDGATLRSSDQATWITPLPKLGNMPPGMRGERLAVDNNNIIYFAPSGKGLWRS 200
DB 128 GNGAVILRGDRGASQKQKTDLPKLGNNPPGMRGERLAVDNNNIIYFAPSGKGLWRS 187
QY 201 TDSGATWSQMTNFPDVGTYIANPTDTTCYQSDIQGVWVWAFDKSSSLGQASKTIFVGVA 260
DB 188 TDSGASNSQVTDFFNVGVYVQDAUTDSYASDNQGIWVTFDESGSGSRTIVVGVA 247
QY 261 DPNPNFVWRDGGATWQAVPGATFIPKHGVFDPVNNHVLXIATSGTGPYDSSGSDVMK 320
DB 248 DXDNSVYRSTAGATWSRLAGQPTGHLAKHGVLDAAANGCLYLAYSCKGPDYDGGKGLWR 307

QY 321 PSVTSCTWTRISVPSTDTANDYFGYSGLTIDROHPNTIMVATQISWMPDIIIFRSTDCG 380
DB 308 YTTKTGTNTNISFVABADT---YGFSGLTVDROHFGTVMATAYSSWMPDTQLFRSTDSG 364
QY 381 ATWTRIDWTSVPNSRLRYVLDISAEPMWLTFFGQPNPVPSPKLGWDEAMADPPNSDR 440
DB 365 GTWTXANDYTSVPNSRNRFTMDVSSPMLTGANPAPPEQTPKLGNWTSLEIDPFD SAR 424
QY 441 MLYGTGATLYATNDLTWDSGGQIHIAPVMVGLTEAVNDLISPPS--GAPLISALGDLGG 499
DB 425 WMYGTGATVYGTNDLTNWDSSGQFTIKPMARGLEETAVNDLASPPSGGAQLFSAIGDGG 484
QY 500 FTHADVAVPSTIFSPVETTTGTSVDYAEINPSIIVRAGSFDPSOPNDRHVAFTSDGGK 559
DB 485 FRHTDLTTVPVLSMTYSPNFTTSLDYAETDGTVVRVGNLD--SGP---HVAFTSDNGA 539
QY 560 NMFQSEPCGGVTTGGTVAASADGSRFVWAPGDPQPVVAVGFGNSWAASQGVANAOIR 619
DB 540 NMFAGADPSGVSGGTVAASDGSRFVWSPAGTG---VQYTTGFTGWSASAGLPAGAVE 597
QY 620 SDRVNPKTIFYALNSGTFFYRSTDDGGVTFQPVAA--GLPSSGAVGMFHAVPKGEGDLMLAAS 678
DB 598 SDRVDPKTFYGFKSGRFYVSSDGGATFTASAATGLPSGDS--VRFKALPCTKGDIMLAGG 655
QY 679 S-----GLYHSTNGSSWSAITGVSSAVNVGFKSAPGSYPVAVFVGTGGVYGVYRSD 733
DB 656 ASDGAYGLWHSTDDGGAFTKLATVDQADTIGFKAATGASQYTLTYSAKIGGVYRGI FRST 715
QY 734 DCGTTWVLINDDQHQYGNWGOAITGDHANLRVYIGTNGRGIVYGDIGGAPSGSPSPSVS 793
DB 716 DKGASWTRVNDDAHOMGWTGAATGDPVYGVYVTVSTNGRGIVYGTAGSDGG----- 769
QY 794 PSASPSLSPSPPSSPSPPSSPSPPSSPSPPSSPSPPSSPSPPSSPSPPSSPSPPSSPS 853
DB 770 ----- 769
QY 854 SSPSSPSPTPSSSPVSGGVKQYKNDNAPGDNQIKPGLQVNVNTGSSSYDLSTVTVRYW 913
DB 770 ---GTEPAPT-----GACTVYRITNQWSGGFQ--ADVQLANTGCTANDGWSLG---W 814
QY 914 FTRDGSSTLVYNCDMAAIGCG-----NTRASFGS 943
DB 815 SPGDGOEVTQLWNASYAQAGSGVTAANLAWNGRVAAGS 852

RESULT 8
US-09-917-376-7
; Sequence 7, Application US/09917376
; Publication No. US20040038334A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
; FILE REFERENCE: 40197.4US01
; CURRENT APPLICATION NUMBER: US/09/917,376
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Aspergillus aculeatus
US-09-917-376-7

Query Match 32.7%; Score 1680; DB 3; Length 726;
Best Local Similarity 46.2%; Pred. No. 4e-86;
Matches 346; Conservative 113; Mismatches 258; Indels 32; Gaps 17;
QY 47 ATTPQVYNSVAVI--GGGGFVDGIVNEGAPGILYVRTDIGGMYRWDAAANGRWITPLLDWVG 105


```
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 838
; TYPE: PRT
; ORGANISM: Trichoderma reesei
US-10-420-191-2

Query Match      31.7%; Score 1625.5; DB 4; Length 838;
Best Local Similarity 41.7%; Pred. No. 5.5e-83;
Matches 350; Conservative 131; Mismatches 296; Indels 63; Gaps 23;

Qy 25 SFVAALGVLPALITASPAAHAATQPYTWSNVAI-GGGGFDGIVFNEGAPGILYVTD 83
Db 4 SRVLALVLGAV-----IPAHAA-----PSWKVNLGGGGFVPGIIFHPKTKGVAYARTD 53

Qy 84 IGGMYRDAANGRIPLLDWV-----GNNWYGVVSVIAADPINTNKVAAVGMVYNSWD 139
Db 54 IGGLYRLUN-ADDSMTAVTDGLADNAGHNW---GIDAVALDPQDDQKVYAAVGMVYNSWD 109

Qy 140 PNDGAILRSSDQGATWQITPLPFKLGNNMPCRGMRGLAVDPNNNDNIIYFGAPSGKGLWR 199
Db 110 PSNGAILRSSDQGATWSTNLPFKVGNMPCRGMRGLAVDPANSNIIYFGARSGNGLWK 169

Qy 200 STDGATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSSSLGQASKTIFVGV 259
Db 170 STDGVTFSKVSFTATCTYIPDPSDSNGYNSDKQLMWVTFDSTSTTGCATSRIFVGT 229

Qy 260 ADP-NNPVFWRDGGATWQAVPGAPTGFPIPHKGVDPDNNHLYIATNTGGPYGSSGDV 318
Db 230 ADNTASVYVSTNAGTWSAVPGQFGKYPHAKLQPAEKALYIYSDGTGYDGTGLGSV 289

Qy 319 WKFSVTSWTIRISVPSTDTANDYFGYSGLTIDROHENTIMVATQISWPDITLIERSTD 378
Db 290 WRIDAGTWDITIPVSGSDL---YFGGGLDLQKPGTLVVASLNSWPDQALFRSTD 346

Qy 379 CGATWTRIDWTSYPNRSLRYLDIAEPWLTFG-VQPNPPVPS-----PKLGWMDMAAI 433
Db 347 SGTWSPINAWASYPTETYYISITPPAPWIKNNFIDVTSESPDGLIKRLGWMIESLEI 406

Qy 434 DPFNSDRMLYGTGATLYATNDLTWKDSGGQTHIAPMWKLEETAVNDLISPSPGAPLISA 493
Db 407 DPTDSNHLWYGTGMTIFGCHDLTNWDRHNSIQSLADGIEEFSVQDLASAPGSELLAA 466

Qy 494 LGDLGGFTHA---DVTAVPSTIFTPVFTTGTSDVYAEINPSIIVRAGSPDSSQPNDRH 550
Db 467 VGDNGFTFASRNDLGTSPQVWATPTWATSTSDYAGNSVKSVVRVGNVNTAGTQQ----- 521

Qy 551 VAFSTDGKNWFQSGEPGVTTGGTVAASADGRFVWAPGQPVVYAVFGNSWAAASQ 610
Db 522 VAISDGGATWSIDYAADTSMNGGTVAADGDTILWSTASSG---VQRSQFGSPFASVS 578

Qy 611 GVPANAQIRSDRVPNPKTFYALNSGTFFRSTDPGVTFFQVPAAGLSPSSGAVGM---FHAVPG 668
Db 579 SLPAGAVIASDKTNSVFYAGSGTFFYVSDTGSF---TRG-PKLSAGTIRIDIAAHT 634

Qy 669 KEGDLWLAAASGLYHSTNGSSWAI-TGVSSAVNVGFKSAPGSSYPVAVFVGTIGCVT 727
Db 635 TAGTLYVSTDVGIIFRSTDSGTTCQVSTALTNTQIALGVGS-GSNW-NLYAFGT---GPS 690

Qy 728 GA--YRSDCCGTWVLINDDQHGNGWQOATIGDHANLRRVYIGTNGRGIYV---GDIGGA 783
Db 691 GARLYASDGSASWTDIGSGQFGSISDSTKVAGSGTAGQVYVGTNGRGVYAGTGVGG 750

Qy 784 PSGSPSPSVSPSASPLSPSPSPSSSPSPSPSSSPSPSPSPSPSPSPSPSPSPSPSPSP 843
Db 751 TGGT-----SSSTKQSSSTSSASSTTLSSSVVSTRASTVTSRSTSSAAGTGS 801

; GENERAL INFORMATION:
; APPLICANT: Dunn-Coleman, Nigel
; APPLICANT: Goedegebuur, Frits
; APPLICANT: Ward, Michael
; APPLICANT: Yao, Jian
; TITLE OF INVENTION: BGVI Endoglucanase and Nucleic Acids
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: GC698
; CURRENT APPLICATION NUMBER: US/10/026,994
; CURRENT FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Trichoderma reesei
US-10-026-994-2

Query Match      31.2%; Score 1604; DB 4; Length 818;
Best Local Similarity 41.8%; Pred. No. 8.7e-82;
Matches 340; Conservative 128; Mismatches 291; Indels 54; Gaps 21;

Qy 52 YTWSNVAI-GGGGFDGIVFNEGAPGILYVTDIGGMYRDAANGRIPLLDWV-----GW 106
Db 2 PSWKVNLGGGGFVPGIIFHPKTKGVAYARTDIGLYRLN-ADDSMTAVTDGIADNAGW 60

Qy 107 NNGWYGVVSVIAADPINTNKVAAVGMVYNSWDPNDAIIRSSDQGATWQITPLPKLGG 166
Db 61 HNW---GIDAVALDPQDDQKVYAAVGMVYNSWDPSNGAIIRSSDRGATWSTNLPKVG 117

Qy 167 NMPGRGMRGLAVDPNNNDNIIYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDT 226
Db 118 NMPGRGMRGLAVDPANSNIIYFGARSGNGLMKSTDGVTFSKVSFTATGTYIPDPSDS 177

Qy 227 TGYQSDIQGVVWVAFDKSSSLGQASKTIFVGVADP-NNPVFWRDGGATWQAVPGAPT 285
Db 178 NGYNSDKQLMWVTFDSTSTTGCATSRIFVGTADNTASVYVSTNAGTWSAVPGQPK 237

Qy 286 FIPHKGVDPDNNHLYIATNTGGPYDGSQGVWKEVTSCTWTRISVPSTDTANDYFG 345
Db 238 YPHKAKLQPAEKALYIYSDGTGYDGLSVWRYDIAGGTWKTITPVSGSDL---YFG 294

Qy 346 YSGLTIDROHENTIMVATQISWPDITLIERSTDGATWTRIDWTSYPNRSLRYLDISA 405
Db 295 FGGGLDLQKPGTLVVASLNSWPDQALFRSTDSGTTNSPIWAWASYPTETYYISSTPK 354

Qy 406 EPWLTFG-VQPNPPVPS-----PKLGWMDMAAIDPFNSDRMLYGTGATLYATNDLTWDS 460
Db 355 APWIKNNFIDVTSESPDGLIKRLGWMIESLEIDPTDSNHLWYGTGMTIFGCHDLTNWDT 414

Qy 461 GGQTHIAPMWKLEETAVNDLISPSPGAPLISALDGLGFTHA---DVTAVPSTIFTPV 517
Db 415 RHNSIQSLADGIEEFSVQDLASAPGSELLAAVGDNGFTFASRNDLGTSPQVWATPT 474

Qy 518 FTTGTSDVYAEINPSIIVRAGSPDSSQPNDRHAFSTDGKNWFQGSBEQVTTGGTVA 577
Db 475 WATSTSDYAGNSVKSVVRVGN-----TAGTQVALSSDGGATWSIDYAADTSMNGGTVA 528

Qy 578 ASADGRFVWAPGDPQPVVYAVFGNSWAAASQGVPAQAIRSDRVPNPKTFYALSNGTFY 637
Db 529 YSADGDTILWSTASSG---VQRSQFGSPFASVSLPAGAVIASDKTNSVFYAGSGSTFY 585

Qy 638 RSTDGVTFFQVPAAGLSPSSGAVGM---FHAVPGKEGDLWLAAASGLYHSTNGSSWAI- 694
Db 586 VSKDTGSF---TRG-PKLSAGTIRIDIAAHTPTAGTLYVSTDVGIIFRSTDSGTTCQVS 641

Qy 695 TGVSSAVNVGFKSAPGSSYPVAVFVGTIGCVTGA---YRSDCCGTWVLINDDQHGNGW 752
Db 642 TALNTYQIALGVGS-GSNW-NLYAFGT---GPSGARLYASGDSGASWTDIQSGQFGSID 697

Qy 753 QOAITGDHANLRRVYIGTNGRGIYV---GDIGAPSGSPSPSPSPSPSPSPSPSPSPSP 810
Db 698 STKVAGSGTAGQVYVGTNGRGVYAGTGVGGTGT-----SSSTKQSSSTSSA 748
```

RESULT 11

US-10-026-994-2

; Sequence 2, Application US/10026994

; Publication No. US20030113732A1

Db 338 LDRDPGALDSIYLSTDAAGTAKWDTVTQLSSPSNLEGNWGHPTNAARY-KDGTTPVWLDNF 396
Qy 413 VQP-----NPPVSP---KLGMDEMAIDPPNSDRMLYGTGATLYATNLTGK----WDS 460
Db 397 NGPOMGGYGAPHGTPGLTKFGWMSAVLIDPPNPEHLMYGTGATLWATDLSRVEKDW-- 454
Qy 461 GQOIHIAP---MVKGLEETAANDLISPPSGAPLISALGLDGGFTHADVTAVPSTFTSP 516
Db 455 -----ASWYLQIDGIEENALISPPSGAALLSGIDISGMKHDLLTK-PQKMFAG 507
Qy 517 VFTTGTSDYAEALNPSIIRAGSPDPSSQPNDRHVAFTSDGKNW--FGSEPPGVTT-- 572
Db 508 QFSNLDSDIDAAGNFPNVVRAGSSGHEYDSACARGAYATDGDATWTFPTCPGPNASHY 567
Qy 573 -GGTVAASADGSRFVWAP--GDPGQPVVYAVFGNSMAASQGVPA-----NAQIRSDRV 623
Db 568 QGSTIAVDASGSIWSTKLDEQASGPMWYSHDYGKTWS---VPAGDLKAQTANVLSDKV 623
Qy 624 NPKTFYALNSGTFFRSTGGVTFQVVAAGLPSSGAVGVMFHAPGKEDLWL-AASSGLY 682
Db 624 QDGTFYATDGGKFFVSTDDGKSAAAGAGLVT--GTSMLPAVNPWVAGDVWVPEGLF 681
Qy 683 HSTNGGSSWSAI-TGVSAVNVGFKS-----APGSSYPAVVFVGT--IGGVTGAYRSDC 735
Db 682 HSTDFGASFTRVGTANATLVSAGAPKSDGKKASAPSAVFWGDKPGSDIGLYRSDN 741
Qy 736 GTTWLINDDQHQGNWQQAITGDHANLRVYVGTNGRGIVYGD 780
Db 742 GSTWRVNDQHNYSG-PTMIEADPKVYGRVYLTNGRGIVYADL 785

RESULT 14

US-10-395-241-14
; Sequence 14, Application US/10395241
; Publication No. US20040038367A1
; GENERAL INFORMATION:
; APPLICANT: YAOI, Katsuro
; APPLICANT: MITSUISHI, Yasushi
; TITLE OF INVENTION: NOVEL XYLOGLUCAN OLIGOSACCHARIDE-DEGRADING ENZYME, POLYNUCLEOTIDE
; TITLE OF INVENTION: ENCODING THE SAME, AND METHOD OF PREPARING THE ENZYME
; FILE REFERENCE: Q73756
; CURRENT APPLICATION NUMBER: US/10/395,241
; CURRENT FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: JP 2002-83433
; PRIOR FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 14
; LENGTH: 789
; TYPE: PRT
; ORGANISM: Geotrichum sp. M128
US-10-395-241-14

Query Match 21.8%; Score 1117; DB 4; Length 789;
Best Local Similarity 35.0%; Pred. No. 2e-54;
Matches 281; Conservative 124; Mismatches 279; Indels 120; Gaps 31;

Qy 52 YWNSVAIGGGGFDVGVNFGAPGILYVRTDIGMYRWDAAANGRWIPLLDVWGNWNGY 111
Db 4 YEFKNVAIGGGGYITGIVAHPTKDLLYARTDIGAYRWDAGTSKWIPLNDFIEAQDMNI 63
Qy 112 NGVYSIAADPTNKNVAAVGMYT-NSWDPNDGALLRSSDQATWQITPLPKLGNWPG 170
Db 64 MGTESIALDPNPNRLYLQAQRYGVDEW-----AAPVYSEDRQSFTIYESPPFMGANDMG 119
Qy 171 RGMGERLAVDPNNNILYFGAPSGKGLWRSDSGATWSQMTNFPDVGTYANPTDTTCYQ 230
Db 120 RNNGERLAVNPNFNSNEVWMTGT-EGIKSSDRAXTWTNVTSIDPAF-----TNGIGYT 172
Qy 231 SDIQGVVWVAFDKSSSSLGQASKQATIFVGVADPNPNFWSRGGATWQAVPGAPTGF 287
Db 173 S-----VIEDP-----ERNGTIYASATAPQG-MYVTHDGGVSWEPVAGQPSSWLNRT 218

Qy 288 -----PHKGVFDPVNVHLXIATSNCTGPGYDGGSDYVWKFVSVTSGTWTRI- 331
Db 219 TGAFPPDKPASPAPQMKVALTP--NFLYVYADYFGPMGVTFGEVWRQNRRTSGAWDDIT 276
Qy 332 -----SPVPSDTANDYFGYGLTIDROHPNITIMVATQISWMPDTI---IFRSTDCGAT 382
Db 277 PRVGNSSPAPYNNQTFPAGGFCGLSDVATNPRLVIT-LDRDPGALDSIYILSTDAGAT 335
Qy 383 WTRIWDWTS-----YPNRSRLRYVLDISABPWITFGVQP-----NPPVPS- 425
Db 336 WKDVTQLSSPSNLEGNWGHPTNAARY-KDGTTPVWLDNFNGPQWGGYGAPHGTPGLTKFG 394
Qy 426 WNDDEAMADPPNSDRMLYGTGATLYATNLTGK----WDSGQOIHIAP---MVKGLEETA 477
Db 395 WMSAVLIDPPNPEHLMYGTGATLWATDLSRVEKDW-----APSWYLQIDGIEENA 446
Qy 478 VNDLISPPSGAPLISALGLDGGFTHADVTAVPSTFTSPVFTTGTSDYAEALNPSIIVRA 537
Db 447 ILSLSPKSGAALLSGIDISGMKHDLLTK-PQKMFAGPQFSNLDSDIDAAGNFPNVVRA 505
Qy 538 GSFDPPSSQPNDRHVAFTSDGKNW--FGSEPPGVTT--GGTVAASADGSRFVWAP--G 590
Db 506 SSSGHEYDSACARGAYATDGDATWTFPTCPGPNASHYQGSTIAVDASGSIWSTKLD 565
Qy 591 DPGQPVVYAVFGNSMAASQGVPA-----NAQIRSDRVNPKTFYALNSGTFFRSTDGV 644
Db 566 EQASGPMWYSHDYGKTWS---VPAGDLKAQTANVLSDKVQDGTFYATDGGKFFVSTDDGK 621
Qy 645 TFQPVAAAGLPSSGAVGVMFHAPGKEDLWL-AASSGLYHSTNGGSSWSAI-TGVSAVN 702
Db 622 SYAAKGAGLVT--GTSMLPAVNPWVAGDVWVPEGLFHTSDTDFGASFTRVGTANATLVS 679
Qy 703 VGFGKS-----APGSSYPAVVFVGT--IGGVTGAYRSDCCTTWLINDDQHQGNWQAI 756
Db 680 VGAPKSDGKKASAPSAVFWGDKPGSDIGLYRSDNNGSTWRVNDQHNYSG-PTMI 738
Qy 757 TGDHANLRVYVGTNGRGIVYGD 780
Db 739 EADPKVYGRVYLTNGRGIVYADL 762

RESULT 15

US-10-395-241-18
; Sequence 18, Application US/10395241
; Publication No. US20040038367A1
; GENERAL INFORMATION:
; APPLICANT: YAOI, Katsuro
; APPLICANT: MITSUISHI, Yasushi
; TITLE OF INVENTION: NOVEL XYLOGLUCAN OLIGOSACCHARIDE-DEGRADING ENZYME, POLYNUCLEOTIDE
; TITLE OF INVENTION: ENCODING THE SAME, AND METHOD OF PREPARING THE ENZYME
; FILE REFERENCE: Q73756
; CURRENT APPLICATION NUMBER: US/10/395,241
; CURRENT FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: JP 2002-83433
; PRIOR FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 18
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Xyloglucan Oligosaccharide-Degrading Enzyme with Histidine Tag
US-10-395-241-18

Query Match 21.7%; Score 1114; DB 4; Length 826;
Best Local Similarity 35.0%; Pred. No. 3.1e-54;
Matches 281; Conservative 123; Mismatches 280; Indels 120; Gaps 31;

Qy 52 YWNSVAIGGGGFDVGVNFGAPGILYVRTDIGMYRWDAAANGRWIPLLDVWGNWNGY 111
Db 5 YEFKNVAIGGGGYITGIVAHPTKDLLYARTDIGAYRWDAGTSKWIPLNDFIEAQDMNI 64

Search completed: March 2, 2006, 14:26:50
Job time : 188.353 secs

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).

Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).


```

; APPLICANT: XU, JIANFENG
; TITLE OF INVENTION: METHODS OF PRODUCING PEPTIDES/PROTEINS IN PLANTS AND
; TITLE OF INVENTION: PEPTIDES/PROTEINS PRODUCED THEREBY
; FILE REFERENCE: 27211/04130
; CURRENT APPLICATION NUMBER: US/11/036,256
; CURRENT FILING DATE: 2005-01-14
; PRIOR APPLICATION NUMBER: 60/602,562
; PRIOR FILING DATE: 2004-08-18
; PRIOR APPLICATION NUMBER: 60/582,027
; PRIOR FILING DATE: 2004-06-22
; PRIOR APPLICATION NUMBER: 60/536,486
; PRIOR FILING DATE: 2004-01-14
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: Patent In Ver. 3.3
; SEQ ID NO 33
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: amino acid construct
US-11-036-256-33

Query Match 4.1%; Score 209; DB 7; Length 346;
Best Local Similarity 43.4%; Pred. No. 4.9e-05;
Matches 49; Conservative 16; Mismatches 22; Indels 26; Gaps 6;

QY 783 APSGSPSPVSP-SASPSLSPSPPSS-----SPSPSPSP-----SSSPSSSPS 825
Db 148 SPSPTTTPPLGPHSPPTLSPSTPTPPPGPHSPPPPLSPSPSTPTTTPPLGPHSPPTLSPS 207

QY 826 PSPSPSPSPSRGSPSPSASPS-----SSPSPPSSSPSPSPSPSPSPVSG 871
Db 208 PFTTTPPGP-HGPPPLSPSPFTTTPPLGPHSPPTLSPSPSTTTP-PGVPVSG 258

RESULT 7
US-11-036-256-21
; Sequence 21, Application US/11036256
; Publication No. US20060026719A1
; GENERAL INFORMATION:
; APPLICANT: XU, JIANFENG
; TITLE OF INVENTION: METHODS OF PRODUCING PEPTIDES/PROTEINS IN PLANTS AND
; TITLE OF INVENTION: PEPTIDES/PROTEINS PRODUCED THEREBY
; FILE REFERENCE: 27211/04130
; CURRENT APPLICATION NUMBER: US/11/036,256
; CURRENT FILING DATE: 2005-01-14
; PRIOR APPLICATION NUMBER: 60/602,562
; PRIOR FILING DATE: 2004-08-18
; PRIOR APPLICATION NUMBER: 60/582,027
; PRIOR FILING DATE: 2004-06-22
; PRIOR APPLICATION NUMBER: 60/536,486
; PRIOR FILING DATE: 2004-01-14
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: Patent In Ver. 3.3
; SEQ ID NO 21
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: amino acid construct
US-11-036-256-21

Query Match 4.0%; Score 206; DB 7; Length 183;
Best Local Similarity 44.2%; Pred. No. 3.5e-05;
Matches 50; Conservative 16; Mismatches 23; Indels 24; Gaps 7;

QY 783 APSGSPSPVSP-SASPSLSPSPPSS-----SPSPSPSP-----SSSPSSSPS 825
Db 72 SPSPTTTPPLGPHSPPTLSPSTPTPPPGPHSPPPPLSPSPSTTTPPLGPHSPPTLSPS 131

```

Qy 826 PSPSPSPSRSPSPASPSPPSP--SPSSPSS-SPSPTPSSSPVSGVKV 875
 Db 132 PTPTPPPGP-HSPPPPLSPSPPTPPLGPHSPPTLSSTPTPPP---GVPV 180

```

RESULT 8
US-10-485-517-351
; Sequence 351, Application US/10485517
; Publication No. US2005025629A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynexus Incorporated
; APPLICANT: Foster, Simon
; APPLICANT: Mond, James
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P100629W0
; CURRENT APPLICATION NUMBER: US/10/485,517
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118925.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 351
; LENGTH: 743
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-485-517-351

```

Query Match	4.0%;	Score 205;	DB 6;	Length 743;
Best Local Similarity	23.4%;	Pred. No. 0.00019;		
Matches	108;	Conservative	43;	Mismatches 179; Indels 132; Gaps 19;

Qy	507	AVPSTTFTSPVFTT	----	GTSVDYAE	LNPSII	VRAGS	FD	SSQPNDR	RHVA	FTDGGK	---	559
				:	:	:	:	:	:	:	:	
Db	60	AVRSLAAVEP	VNAADAKG	TNNVDK	VTASNF	KLEKTT	FD	PNQSGNT	FMAANF	TVT	DKVKS	119
				:	:	:	:	:	:	:	:	
Qy	560	-NWFQSGSE	PGVTTGGT	VAAASAD	SGRFV	MAP	---	-GDPQ	QPVVY	AV	----	603
				:	:	:	:	:	:	:	:	
Db	120	GXYFTAKL	PDLSL	TGNGD	VYDSNN	TWPIADI	IKS	TNGD	VVAKA	TYDIL	TKTYFV	FTDY
				:	:	:	:	:	:	:	:	
Qy	604	NSWAASQ	-----	VPANAQR	SRVNP	KTFV	ALNSGT	FY	STDCG	GV	---	645
				:	:	:	:	:	:	:	:	
Db	180	NNKENING	QFSLP	FTDR	AKAPK	SGTY	DANIN	ADE	MFN	KIY	N	----
				:	:	:	:	:	:	:	:	
Qy	646	FOPVAAGL	PSGAV	GVNFH	AVP	KEGDL	WLAAS	GLY	HS	T	----	697
				:	:	:	:	:	:	:	:	
Db	234	-KPNGANI	-SSQII	IGVD	-----	TASQ	NTYKQ	TFV	NP	KQ	RV	LGNTWV
				:	:	:	:	:	:	:	:	
Qy	698	SSAVNV	GFKSA	PGSSY	PAV	VVVT	GTG	TV	GRSD	-----	---	733
				:	:	:	:	:	:	:	:	
Db	280	QDKIES	SGKVS	ATDK	LRI	FVND	T	SKLS	D	SYAD	P	NSLKE
				:	:	:	:	:	:	:	:	
Qy	734	-----	DCG	TW	LIND	QHOY	GNW	QAIT	G	DHAN	LRR	VY
				:	:	:	:	:	:	:	:	
Db	340	ASIK	FGDIT	KTYV	VL	VEG	--	HYD	NTGK	-----	NLKTQV	IQENV
				:	:	:	:	:	:	:	:	
Qy	773	RGIV	-Y	GDI	G	AGP	SG	SPS	-	VSP	SA	SPSL
				:	:	:	:	:	:	:	:	
Db	391	ENVVR	G	---	GG	SAD	G	S	AVN	KP	DT	CP
				:	:	:	:	:	:	:	:	
Qy	831	SP	-----	SPSR	SP	SA	SP	SP	SP	SP	SP	SP
				:	:	:	:	:	:	:	:	
Db	449	DPD	S	D	S	D	S	D	S	D	S	D
				:	:	:	:	:	:	:	:	

RESULT 9
US-10-485-517-200
; Sequence 200, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield

```

? APPLICANT: Biosynexus Incorporated
? APPLICANT: Foster, Simon
? APPLICANT: Mond, James
? TITLE OF INVENTION: Antigenic Polypeptides
? FILE REFERENCE: P100629WO
? CURRENT APPLICATION NUMBER: US/10/485,517
? CURRENT FILING DATE: 2004-02-02
? PRIOR APPLICATION NUMBER: GB 0118825.9
? PRIOR FILING DATE: 2001-08-02
? PRIOR APPLICATION NUMBER: GB 0200349.9
? PRIOR FILING DATE: 2002-01-09
? NUMBER OF SEQ ID NOS: 424
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 200
? LENGTH: 877
? TYPE: PRT
? ORGANISM: Staphylococcus aureus
? US-10-485-517-200

```

```

Query Match      4.0%; Score 205; DB 6; Length 877;
Best Local Similarity 23.4%; Pred. No. 0.00024;
Matches 108; Conservative 43; Mismatches 179; Indels 132; Gaps 19;

Qy 507 AVPSSTTFTSPVFTT---GTSVDVYAEINPSIIIVRAGSFDESSQPNDRHVAFSTDGK---559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 194 AVRSLAAEVPVNAADAKGTNVNDKVYATSNFKLEKTTTFDPNQSGNTPMAANFTVTDVKVS 253

Qy 560 -NPFQSEPGCVTTGGTVAASADGSRFVWAP-----GDCQPVVYAV-----GFG 603
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 254 GDYFTAKLPDSLTCNGDVYDSNNNTMPIADIKSTLGDVVAKATYDILTKTYTFVFTDYV 313

Qy 604 NSWAASQG-----VPAQAQIRSDRVNPKTYALSNGTYFYRSTDCGVT 645
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 314 NNKENINGOFSLPFTDRAKAPKSGTYDANINIADEMNNKITYN-----YSSPIAGID 367

```

646	Qy	FQPVAAGLPSSGAVGVWFHAPVPGKEGDLWLAASSGLYHST---	NG-----GSSWSAITGV	697
368	Db	-KPNGANI--SSQIIIGVD-----	TASGQNTYKQTFVNPQRVLGNTWYIKGY	413
698	Qy	SSAVNVFGKSGAPSSYPAVFVVGITGGVTCAYRS-	-----	733
414	Db	QDKIESSGKVSATDTKLRIPEVNDTSKLSUSYYADPNDSNLKEVTDQKRIYYEHPNV	473	
734	Qy	-----DCGTTTWLINDHQHOYGNWGQAITGDHANLRVYI-	-----GTNG	772
474	Db	ASIKFGDITKTYVVLVEG--HYDNTGK-----	NLKTQVIQENVDPVTRNDRYSIFGWNN	524
773	Qy	RGIV-YGDIIGAGPSGSPS-VSPSAPSLSPSPSSSPSPSPSPSPSPSPSPSPSPSP	830	
525	Db	ENVRYG--CGSADGDSAVNPKDTPPGFVDPFSPDPPEPPTPDPEPSPDPPEPSPDP	582	
831	Qy	SP-----SPSRSP	867	
583	Db	DPDSDDSDSDSGSDSDSDSGSDSDSDSGSDSDSDSDSDSDSDSDSDSDSDSDS	624	

```

RESULT 10
US-11-108-172-1065
; Sequence 1065, Application US/11108172
; Publication No. US20050260177A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yuyu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon B.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.

```


QY 765 -----RVXIGTN-GRGIV-----YGDIGAPSGSPSPSPSPSPSPSPSPSS 808
Db 786 PEIALTFGVDPPIHIPINIDAGVVTLOQFSIVAENNIDFTPIIPTINILPTAAITVG 845
QY 809 SPSS 867
Db 846 GPTTSIGITASAGISITIPIDPATSGFGNSTSPSSGFGNSGAGSAGSLNVVAGAS 905
QY 868 PVSGGVKVOYKXNDAPGDNQIKPGLQVNTG-----SSVDLSTVTVRYNFTRDG 919
Db 906 GISG-----YLVNGAL-----GSGVTNVGHTVSGFYNASALDLVTPAFASGLMRD-G 951
QY 920 SSTLVNCDWAAIGCGNIRASFGS 943
Db 952 MGTWTLNLGLANLGN--AGFGN 973

RESULT 14
US-11-087-099-7521
; Sequence 7521, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 7521
; LENGTH: 7465
; TYPE: PRT
; ORGANISM: Magnetococcus sp. MC-1
US-11-087-099-7521

Query Match 3.8%; Score 197.5; DB 7; Length 7465;
Best Local Similarity 18.8%; Pred. No. 0.007;
Matches 211; Conservative 105; Mismatches 461; Indels 346; Gaps 41;

QY 17 VSLAATAS-----PAAAAAGVLPPIATASPAHAATTPQYTWSNVAGGG----- 63
Db 6460 VSTLTGSSASINEDTSVSGTLTVVDADVIDATVTAATSVG-TYGTFSVGSNGVMSYQ 6518

QY 64 -----FVDGIVFNEGAPGILYVRTDIGMY-----RMDAANGRWIPLLDVYGMNW 109
Db 6519 LDNSKAVVOGLTQGOVSESTVHTSDGSSSLTNITGRQDSA----- 6562

QY 110 GYNGVVSIAADPINTNKVAAVGMYT-NSWDPNDGAILRSSDOGA-----TWQ 156
Db 6563 ----VVGSGTGSVTESTLSSGMLSVSDADSGEAAVYASTGTGSGFSLDSSGAWSYQ 6618

QY 157 ITPLPFKLGNNPGRGMGERL---AVDPNNIL----- 187

Db 6619 LNNASQOQALIMQGVSESTVATVDGTASSVVTVVGAQDAQAQGGVQSGVVEDQOL 6678

QY 188 -----YFGAPSGKGLWR-----STDGATWSQNTNPDVGTIANPTDITGY 229
Db 6679 STAGLTISDADAGSLTQVASTDGMVGHFAMDSGAWSS-----YVLN-NEIAAI 6727

QY 230 QSDIQG-VVWVAFDKSSSLQASKTIFVGVADPNNPVFWSRDGG---ATQAVPGAPTQ 285
Db 6728 QOMIAGQTVLESFTVSSSDGQVQVSVTRI-TGSQDNVITGEAAGQVSETNDDQSQMATG 6786

QY 286 FIPHKGVDPNVHVLVYATSNITGPGYDSSGVNKFVS-TGTWTRI-----SPVPS 336
Db 6787 KLNISLDSQQAHFV-----GGPQAGYQ---AFTLQSGDNNTVLDNHNHGVAMAGLT 6836

QY 337 TDANDYGYSLTIDRHPNTIMVATQISHWPTIIFRSTGDNATWTRIDWTSYNPRS 396
Db 6837 GDVVTETTVRSL--DNFAETTV-----TITINGLDEPTTAEIEAARLAEQR 6883

QY 397 LRYVLD-----ISAEPLWTFGVQPNPVPKLGWMDMAIDPFNSDRMLYGTGA 447

Db 6884 LQELQASNDLLQGAIGAEGNAGGTATPPPPNPNPADGAPG-----LLGAGG 6931
QY 448 TLYATNDLTKWDSGGQIHIHPMVKGLETAENVDLISPPSGAPLISALGDLGGF--THADV 505
Db 6932 A--AAGGATGGTGGGTAGTGGGGLGGGFG-----SCAPNTSAGLGTGGGTATGG 6983
QY 506 TAVPSTIFTSVFTTCTSDVYAEALNPSIIVRAGSPDPSPQPNDRHVAFSTDGKKNWFQGS 565
Db 6984 TVTGGTGGTGAATGGTA-----TGGTATGGTGGTGGT 7016

QY 566 EPGGVTTGTV-----AASADGRFVWAPDPGPQPVVAVG----- 601
Db 7017 TTGGTATGGTATGGTGAAGGTGGAAGGA-----ATGEAGGTATGGTGGTGGTATGGTAT 7072

QY 602 -----FGNSWAASQGVPANAIQRSDRVNPKTFYALSNGTFYRSTDG--GVTFQPVAAGL 653
Db 7073 CGTGGAAGGTGGAAGGAATGE-AGGTATGGTGGTATGGTATGGTGAAGGTGGAAGGA 7131

QY 654 PSSGAVGVWFMHAVPG-KEGDLMLAASSGLYHSTNGSSWSAITG-VSSAVNVGFGKSAPG 711
Db 7132 ATGEAGGAATGGTGGTATGGTATGGTGAAGGTGGAAGGAATGEAGGAATGGTGAAG 7191

QY 712 SSYPAVFVVGTIGGVTGAYRSDCGTTWVLINDDQHQYGNWGOAITGDHANLRVIGTN 771
Db 7192 GTGGA--ATGEAGGAATGGTADAGDT-----GGAATGGDAN-----CGAA 7230

QY 772 GRGIVGIDIGGAPSGSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSS 824
Db 7231 GPGVVDGGAAGGPGADGGAAGGPGADGGAAGGPGANGGAAGGPGVDDGGAAGGPGVDGGA 7290

QY 825 -----SP 839
Db 7291 GPGADGGAAGGPGAVPEGGDPAGANGQGAAGQDGTNDTNDTDEAPADGEAPAEAGPPA 7350

QY 840 PSASPS 899
Db 7351 DGEAPAEQVQPADGEAPAEQVQPADGEAPAEQVQPADGEAPAEQVQPADGEAPAEQVQ 7406

QY 900 SSSVDLSTVTVRYWFTRDGSGSTLVVNCDDWAAIGCGNIRASFG 942
Db 7407 SGTGEAI-----AGSAAIA-----AAWGGMGRIQWG 7433

RESULT 15
US-11-036-256-160
; Sequence 160, Application US/11036256
; Publication No. US20060026719A1
; GENERAL INFORMATION:
; APPLICANT: KIELISZEWSKI, MARCIA
; APPLICANT: XU, JIANFENG
; TITLE OF INVENTION: METHODS OF PRODUCING PEPTIDES/PROTEINS IN PLANTS AND
; TITLE OF INVENTION: PEPTIDES/PROTEINS PRODUCED THEREBY
; FILE REFERENCE: 27211/04130
; CURRENT APPLICATION NUMBER: US/11/036,256
; CURRENT FILING DATE: 2005-01-14
; PRIOR APPLICATION NUMBER: 60/602,562
; PRIOR FILING DATE: 2004-08-18
; PRIOR APPLICATION NUMBER: 60/582,027
; PRIOR FILING DATE: 2004-06-22
; PRIOR APPLICATION NUMBER: 60/536,486
; PRIOR FILING DATE: 2004-01-14
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: Patent in Ver. 3.3
; SEQ ID NO 160
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: amino acid motif
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (1)..(1)

1	OTHER INFORMATION:	Lys, Ser, Ala, Thr, Gly, or Val
2	FEATURE:	
3	NAME/KEY:	MOD RES
4	LOCATION:	(3) .. (3)
5	OTHER INFORMATION:	Lys, Ser, Ala, Thr, Gly, or Val
6	FEATURE:	
7	NAME/KEY:	MOD RES
8	LOCATION:	(5) .. (5)
9	OTHER INFORMATION:	Lys, Ser, Ala, Thr, Gly, or Val
10	FEATURE:	
11	NAME/KEY:	MOD RES
12	LOCATION:	(7) .. (7)
13	OTHER INFORMATION:	Lys, Ser, Ala, Thr, Gly, or Val
14	FEATURE:	
15	NAME/KEY:	MOD RES
16	LOCATION:	(9) .. (9)
17	OTHER INFORMATION:	Lys, Ser, Ala, Thr, Gly, or Val
18	FEATURE:	
19	NAME/KEY:	MOD RES
20	LOCATION:	(11) .. (11)
21	OTHER INFORMATION:	Lys, Ser, Ala, Thr, Gly, or Val
22	FEATURE:	
23	NAME/KEY:	MOD RES
24	LOCATION:	(13) .. (13)
25	OTHER INFORMATION:	Lys, Ser, Ala, Thr, Gly, or Val
26	FEATURE:	
27	NAME/KEY:	MOD RES
28	LOCATION:	(15) .. (15)
29	OTHER INFORMATION:	Lys, Ser, Ala, Thr, Gly, or Val
30	FEATURE:	
31	NAME/KEY:	MOD RES
32	LOCATION:	(17) .. (17)
33	OTHER INFORMATION:	Lys, Ser, Ala, Thr, Gly, or Val
34	FEATURE:	
35	NAME/KEY:	MOD RES
36	LOCATION:	(19) .. (19)
37	OTHER INFORMATION:	Lys, Ser, Ala, Thr, Gly, or Val
38	FEATURE:	
39	NAME/KEY:	MOD RES
40	LOCATION:	(21) .. (21)
41	OTHER INFORMATION:	Lys, Ser, Ala, Thr, Gly, or Val
42	FEATURE:	
43	NAME/KEY:	MOD RES
44	LOCATION:	(23) .. (23)
45	OTHER INFORMATION:	Lys, Ser, Ala, Thr, Gly, or Val
46	FEATURE:	
47	NAME/KEY:	MOD RES
48	LOCATION:	(25) .. (25)
49	OTHER INFORMATION:	Lys, Ser, Ala, Thr, Gly, or Val
50	FEATURE:	
51	NAME/KEY:	MOD RES
52	LOCATION:	(27) .. (27)
53	OTHER INFORMATION:	Lys, Ser, Ala, Thr, Gly, or Val
54	FEATURE:	
55	NAME/KEY:	MOD RES
56	LOCATION:	(29) .. (29)
57	OTHER INFORMATION:	Lys, Ser, Ala, Thr, Gly, or Val
58	FEATURE:	
59	NAME/KEY:	MOD RES
60	LOCATION:	(31) .. (31)
61	OTHER INFORMATION:	Lys, Ser, Ala, Thr, Gly, or Val
62	FEATURE:	
63	NAME/KEY:	MOD RES
64	LOCATION:	(33) .. (33)
65	OTHER INFORMATION:	Lys, Ser, Ala, Thr, Gly, or Val
66	FEATURE:	
67	NAME/KEY:	MOD RES
68	LOCATION:	(35) .. (35)
69	OTHER INFORMATION:	Lys, Ser, Ala, Thr, Gly, or Val
70	FEATURE:	
71	NAME/KEY:	MOD RES
72	LOCATION:	(37) .. (37)
73	OTHER INFORMATION:	Lys, Ser, Ala, Thr, Gly, or Val

1	FEATURE:	
2	NAME/KEY: MOD RES	
3	LOCATION: (39)..(39)	
4	OTHER INFORMATION: Lys, Ser, Ala, Thr, Gly, or Val	
5	FEATURE:	
6	NAME/KEY: MOD RES	
7	LOCATION: (41)..(41)	
8	OTHER INFORMATION: Lys, Ser, Ala, Thr, Gly, or Val	
9	FEATURE:	
10	NAME/KEY: MOD RES	
11	LOCATION: (43)..(43)	
12	OTHER INFORMATION: Lys, Ser, Ala, Thr, Gly, or Val	
13	FEATURE:	
14	NAME/KEY: MOD RES	
15	LOCATION: (45)..(45)	
16	OTHER INFORMATION: Lys, Ser, Ala, Thr, Gly, or Val	
17	FEATURE:	
18	NAME/KEY: MOD RES	
19	LOCATION: (47)..(47)	
20	OTHER INFORMATION: Lys, Ser, Ala, Thr, Gly, or Val	
21	FEATURE:	
22	NAME/KEY: MOD RES	
23	LOCATION: (49)..(49)	
24	OTHER INFORMATION: Lys, Ser, Ala, Thr, Gly, or Val	
25	FEATURE:	
26	NAME/KEY: MOD RES	
27	LOCATION: (51)..(51)	
28	OTHER INFORMATION: Lys, Ser, Ala, Thr, Gly, or Val	
29	FEATURE:	
30	NAME/KEY: MOD RES	
31	LOCATION: (53)..(53)	
32	OTHER INFORMATION: Lys, Ser, Ala, Thr, Gly, or Val	
33	FEATURE:	
34	NAME/KEY: MOD RES	
35	LOCATION: (55)..(55)	
36	OTHER INFORMATION: Lys, Ser, Ala, Thr, Gly, or Val	
37	FEATURE:	
38	NAME/KEY: MOD RES	
39	LOCATION: (57)..(57)	
40	OTHER INFORMATION: Lys, Ser, Ala, Thr, Gly, or Val	
41	FEATURE:	
42	NAME/KEY: MOD RES	
43	LOCATION: (59)..(59)	
44	OTHER INFORMATION: Lys, Ser, Ala, Thr, Gly, or Val	
45	FEATURE:	
46	NAME/KEY: MOD RES	
47	LOCATION: (61)..(61)	
48	OTHER INFORMATION: Lys, Ser, Ala, Thr, Gly, or Val	
49	FEATURE:	
50	NAME/KEY: MOD RES	
51	LOCATION: (63)..(63)	
52	OTHER INFORMATION: Lys, Ser, Ala, Thr, Gly, or Val	
53	FEATURE:	
54	NAME/KEY: MOD RES	
55	LOCATION: (65)..(65)	
56	OTHER INFORMATION: Lys, Ser, Ala, Thr, Gly, or Val	
57	FEATURE:	
58	NAME/KEY: MOD RES	
59	LOCATION: (67)..(67)	
60	OTHER INFORMATION: Lys, Ser, Ala, Thr, Gly, or Val	
61	FEATURE:	
62	NAME/KEY: MOD RES	
63	LOCATION: (69)..(69)	
64	OTHER INFORMATION: Lys, Ser, Ala, Thr, Gly, or Val	
65	FEATURE:	
66	NAME/KEY: MOD RES	
67	LOCATION: (71)..(71)	
68	OTHER INFORMATION: Lys, Ser, Ala, Thr, Gly, or Val	
69	FEATURE:	
70	NAME/KEY: MOD RES	
71	LOCATION: (73)..(73)	
72	OTHER INFORMATION: Lys, Ser, Ala, Thr, Gly, or Val	
73	FEATURE:	

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 2, 2006, 12:06:58 ; Search time 345.215 Seconds
(without alignments)
4927.737 Million cell updates/sec

Title: US-09-917-376-1
Perfect score: 5134
Sequence: 1 MDRSENRLTWRSRLVSL.....RASFGSNPATPTADTYLQX 957

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q-/abss/ABSSWEB_spool/US0917376/runat 02032006_091454_8101/app_query.fasta_1
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss06h -USER=US0917376 @CGN 1.1 278 @runat 02032006_091454_8101
-NCPU=6 -ICPU=3 -NO_WMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PCUS COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	409	8.0	11707	3	US-09-136-574A-1
2	392.5	7.6	1103	3	US-09-533-559-7511
C 3	358	7.0	19383	3	US-09-949-016-16031
C 4	345	6.7	18596	3	US-09-318-448-11
C 5	345	6.7	18596	3	US-09-577-286-11
C 6	345	6.7	18597	3	US-09-962-665-8
C 7	345	6.7	18597	3	US-09-963-333-8
C 8	345	6.7	18597	3	US-09-962-677-8
C 9	334	6.5	6416	3	US-09-136-574A-2

10	319.5	6.2	8211	3	US-09-252-991A-13656	Sequence 13656, A
11	296.5	5.8	4767	3	US-09-410-551B-28	Sequence 28, Appl
12	296.5	5.8	4767	3	US-09-940-316B-28	Sequence 28, Appl
13	293	5.7	4818	3	US-09-410-551B-32	Sequence 32, Appl
14	293	5.7	4818	3	US-09-940-316B-32	Sequence 32, Appl
C 15	288.5	5.6	4131	3	US-09-252-991A-13773	Sequence 13773, A
C 16	288.5	5.6	4571	3	US-09-410-551B-18	Sequence 18, Appl
17	288.5	5.6	4571	3	US-09-940-316B-18	Sequence 18, Appl
18	288	5.6	288	3	US-09-119-507B-111	Sequence 111, App
19	288	5.6	288	3	US-09-547-693-111	Sequence 111, App
C 20	282.5	5.5	441529	3	US-09-103-840A-1	Sequence 1, Appl
C 21	281.5	5.5	4403765	3	US-09-103-840A-2	Sequence 2, Appl
C 22	280	5.5	3129	3	US-09-252-991A-13873	Sequence 13873, A
23	279	5.4	4547	3	US-09-410-551B-22	Sequence 22, Appl
24	279	5.4	4547	3	US-09-940-316B-22	Sequence 22, Appl
25	273	5.3	4674	3	US-09-410-551B-26	Sequence 26, Appl
26	273	5.3	4674	3	US-09-940-316B-26	Sequence 26, Appl
27	269	5.2	4478	3	US-09-410-551B-16	Sequence 16, Appl
28	269	5.2	4478	3	US-09-940-316B-16	Sequence 16, Appl
C 29	268.5	5.2	31467	3	US-09-949-016-13134	Sequence 13134, A
C 30	268.5	5.2	31868	3	US-09-949-016-11907	Sequence 11907, A
C 31	268	5.2	31391	3	US-09-949-016-14319	Sequence 14319, A
C 32	266	5.2	4188	3	US-09-252-991A-13774	Sequence 13774, A
33	266	5.2	77536	3	US-09-410-551B-1	Sequence 1, Appl
34	266	5.2	77536	3	US-09-940-316B-1	Sequence 1, Appl
C 35	265.5	5.2	77536	3	US-09-410-551B-1	Sequence 1, Appl
C 36	265.5	5.2	77536	3	US-09-940-316B-1	Sequence 1, Appl
37	265	5.2	4737	3	US-09-410-551B-30	Sequence 30, Appl
38	265	5.2	4737	3	US-09-940-316B-30	Sequence 30, Appl
39	264.5	5.2	6210	3	US-10-212-962-1	Sequence 1, Appl
40	263	5.1	30690	3	US-09-914-286-1	Sequence 1, Appl
41	262	5.1	150394	3	US-09-949-016-13042	Sequence 13042, A
C 42	260	5.1	3337	2	US-08-072-610-1	Sequence 1, Appl
C 43	260	5.1	3337	2	US-08-719-822B-1	Sequence 1, Appl
C 44	260	5.1	3337	3	US-09-092-458-1	Sequence 1, Appl
C 45	260	5.1	3337	3	US-08-719-821C-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-09-136-574A-1
; Sequence 1, Application US/09136574A
; Patent No. 6294366

GENERAL INFORMATION:

APPLICANT: Farrington, Graham K.

Anderson, Paige

Gibbs, Moreland

Bergquist, Peter

Daniels, Roy

Morgan, Hugh W.

Williams, Diane P.

TITLE OF INVENTION: Compositions and Methods for Treating Cellulose Containing Fabrics Using Truncated Cellulase Enzyme Compositions

NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Center, P.O. Box 457

CITY: Spring House

STATE: PA

COUNTRY: USA

ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/136,574A

FILING DATE: 19-Aug-1998

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/932,571

FILING DATE: September 19, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Bak, Mary E.

REGISTRATION NUMBER: 31,215

REFERENCE/DOCKET NUMBER: 1997US001/CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-540-9200

TELEFAX: 215-540-5818

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 11707 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-136-574A-1

Alignment Scores:

Pred. No.: 5,45e-11 Length: 11707
Score: 409.00 Matches: 263
Percent Similarity: 32.4% Conservative: 163
Best Local Similarity: 20.0% Mismatches: 386
Query Match: 8.0% Indels: 504
DB: 3 Gaps: 58

US-09-917-376-1 (1-957) x US-09-136-574A-1 (1-11707)

```

QY 14 ArgArgLeuValSerLeuLeuAlaAlaThrAlaSerPheAlaValAlaAlaLeuGly 33
DB 735 AAGAGGGTAATTCATCTT...TCCTTATGTTTATTAACACGCTGTAGGT 791
QY 34 ValLeuProIleAlaIleThrAlaSerProAlaHisAlaAlaThrThrGlnProTyrThr 53
DB 792 ACTTGTG...ATATTTCATCAGGAGCAAGACGACGCA...TATACT 833
QY 54 TrpSerAsnValAlaIleGlyGlyGlyPheValAspGlyIleValPheAsnGluGly 73
DB 834 ...GTTGATTGTAAGGT 848
QY 74 AlaProGlyIleLeuTyr...ValArgThrAspIleGlyGly 86
DB 849 GCTGATACCTTATCTTACTTACTTATGGAATAATCGACATAGCAGTTGACATGGGCAAT 908
QY 87 MetTyrArgTrpAspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrp 106
DB 908 ... 908
QY 107 AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLys 126
DB 909 ...AATAGAAGTTCA 953
QY 127 ValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeu 146
DB 954 ATATGGGATGGAGTTGCAGTT...GACGTTAAA 983
QY 147 ArgSerSerAspGlnGlyAlaThrTrpGlnIleThr... 158
DB 984 AACATTATGACCAATGAACCAATGGTAGTTTCACCGTATGTAAACATAGTACCAG 1043
QY 159 ...ProLeuProPheLysLeu...GlyGlyAsnMetProGlyArg 171
DB 1044 AAGCCGGTTGCAATTTGGTATCTCAGCGGTTTACGACGATCGAAGTGGGTAAAGACTACT 1103
QY 172 GlyMetGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAla 191
DB 1104 CTATAGGTGAGGTTGTGGCTATT...CCAAAT... 1133
QY 192 ProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThr 211
DB 1134 ...TATTGGAAGAAAATT... 1148
QY 212 AsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGlnSer 231

```

```

DB 1149 ...GTTGGTAAATGGACTCAAATATTAGCAATGTACAGGAATTTGTTAATT 1196
QY 232 AspileGlnGlyValValTrpValAlaPheAspLysSerSerSerLeuGlyGlnAla 251
DB 1197 GTAATACACACATTTGTAGAAAGCAAGTAGATTATATGTTGAC... 1241
QY 252 SerLysThrIlePheValGlyValAlaAspProAsn...AsnProValPhe 267
DB 1242 ...TATATCCAAATATGATGATGATATGTTACCTATCAAAATGCAAGTGACA 1289
QY 268 TrpSerArg...AspGlyGlyAlaThr...TrpGlnAla...Val 279
DB 1290 TTTTCAAGTGGATTGAAAGTGGCACTACCGAGGTTGGCAGGCAAGGGAGCGGTGT 1349
QY 280 ProGlyAlaProThrGlyPheIleProHisLysGlyValPheAspProValAsnHisVal 299
DB 1350 ACAGTAAACACAGATAGCGTTGTGCATATATAGTGCAGATATAGT... 1394
QY 300 LeuTyrIleAla...ThrSerAsnThrGlyGly...ProTyrAspGly... 313
DB 1395 TTGTAGCTCAGTGGGAAGAACGTCAAATTTGGCATGTGCACAGATTCGGGTAGATACAATT 1454
QY 314 ...SerSerGlyAspValTrpLysPheSerVal...ThrSerGlyThrTrp 328
DB 1455 TTGGAACAGGCTAAAGTGTATAAAATAAGTGTGTTGGTATTATCAGAACAGTGTGTTCAACT 1514
QY 329 ThrArgIleSerProValProSerThrAspThrAlaAsnAsp...TyrPheGlyTyrSer 347
DB 1515 CAAAATATGTCATTAACTATCCTCAAGAAGATTGTCACAGACTCTTCAACAAGCTATGAA 1574
QY 348 GlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTrp 367
DB 1575 AATCTGATATATACACGGATGTACCGAGTAATACGTGGTTCAGCTGAGTGAAGCTAC 1634
QY 368 TrpProAspThrIleIlePheArgSerThrAspGlyGlyAlaThrTrpThrArgIleTrp 387
DB 1635 ...TCAATTCCTGCTGGTGTACAGTTAGCGAGTGTGTTG 1670
QY 388 AspTrpThrSerTyrProAsnArgSerLeuArgTyr...ValLeu 401
DB 1671 CTTTATGTGAGGCACAAAATGCAAAATTTGGCTTCTCGGTTGATGATTTAAAGATTAT 1730
QY 402 AspileSer...AlaGluPro... 407
DB 1731 GATTATTCCAAGTTGGCTGAACCTGAATGGGAGATACCATCTTTGATAGAAAAGATAGA 1790
QY 407 ... 407
DB 1791 GATTATTTCAAAGTAGGAGTAGCTTTGTCTTACAAAAGCATTGCTCTGTATACAGAAAAG 1850
QY 407 ... 407
DB 1851 AAGATGGTTTTGAAGCATTTTCAATAGTATTACTGCGAGGAACGAAATGAACCATCAGAG 1910
QY 408 ...TrpLeuThrPheGly 412
DB 1911 TTACTTGTGATGAAAATACTTACAACTTTAGCAAGCAGACGAAATTTGTAATTTTGA 1970
QY 413 ValGlnProAsnProProValProSerProLysLeuGlyTrpMetAspGluAla... 430
DB 1971 ACAAGTAACAACATGGCCATCAGAGGTACATACACTGTTTGGCATGAGCAAAACCCGAC 2030
QY 430 ... 430
DB 2031 TGGTTTTTCAAGGACACAAATGGAATACGTTGACGAAGGATGCATTGTGAAGCAGATTA 2090
QY 430 ... 430
DB 2091 AAACAGTATATTTATACGGTAGTGGGAAGATATAAAGGAAGGTTTATGATGATGATGTG 2150
QY 431 ...MetAlaIleAspProPheAsnSerAspArgMetLeu... 442

```

Db 2151 GFAAATGAAGCAATAGATGAAGAGTCAAGGTGATGATTCAGGAGATCTAACTGGTACAAC 2210
Qy 443 -----TyrGlyThrClyAlaThrLeuTyrAla----- 451
Db 2211 ATTTGTAGTCCCGAATATATTGAGAAGGCTTTATATGGGACATGAAGCCGATCCAGAC 2270
Qy 452 -----ThrAsnAspLeuThrLysTrpAspSerGlyGlyGlnIleHisIleAla 467
Db 2271 GCAAAATTTGTTTCAACAGATTACACACAGAAACAGTCAAGAGACAGTATTATTATAC 2330
Qy 468 ProMetValLysGlyLeuGluGluThrAlaValAsnAspLeuLeuSerProProSerGly 487
Db 2331 AACATGATTAGAGTCTCAAGAAAAAGGTGT-----CCATTTCATGGA 2375
Qy 488 AlaProLeuIleSerAlaLeuGlyAspLeuGlyPheThrHisAlaAspValThrAla 507
Db 2376 ATAGGATTCGACAGTCAATA-----AATCTTGATTGGCCCTCGATTAGCGAG 2423
Qy 508 ValProSerThrIle-----PheThrSerProValPheThrGlyThrSerValAsp 525
Db 2424 ATAGAAACACACATAAGATTCTTCAGCTCT-----ATACCTGGATTGGAGATACAC 2474
Qy 526 TyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspProSerSerGln 545
Db 2475 ATTACGGAGCTTGATATAGATTTTATCAGTGGGTTCCGATACCAAGTTACTACACGCCA 2534
Qy 546 ProAsnAsp----- 548
Db 2535 CCAAGAGATCTCCTCATAAACAGCAATGAGATATAGAGTATTATTGATTGTTTAA 2594
Qy 549 -----ArgHisValAlaPhe----- 553
Db 2595 AAGTACAAATGTAATAACAAAGTGAACATCTGGGGAGTGAAGGATGATTACTATCG 2654
Qy 554 ---SerThrAspGlyGlyAsn----- 560
Db 2655 CTGAGTCAAAATTTGGAAAAAGGATTACCCGTTGTTATTGATGAAGAACTATAAATCA 2714
Qy 561 -----TrpPheGlnGlySerGluProGlyGlyValThrGlyThrGlyThrValAla 577
Db 2715 AAATATGCTTTTGGAGCTGATTGAGCCAACTGTGATACCGGCCAACTCAACATTGCCA 2774
Qy 578 Ala-----SerAlaAspGlySerArgPheVal 586
Db 2775 GCACCACCACTATTCAAATACCTACCACTCCACACCAACCCGACACCGACAGTG 2834
Qy 587 TrpAlaProGlyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrp 606
Db 2835 AGTGCACGCCAACACCAACGCA-----CCGACGGCATCACCGGTAGGT---GGCAGTTACTCG 2888
Qy 607 AlaAlaSerGlnGlyValProAla-----AsnAlaGlnIle 618
Db 2889 ACGCGAGTAGAGTTACAGTGCCTGAAGTATGTTGATGGAATGGGAATTTAAGCAGC 2948
Qy 619 ArgSerAspArgValAsnProLysThrPheTyrAlaLeuSerAsn---GlyThr----- 635
Db 2949 CCGAGCATGATTGAATCCTAAG-----ATAAGATAGAGATGTTGGGACGACAGCG 3002
Qy 636 -----PheTyrArgSerThrAspGlyGlyValThr--- 645
Db 3003 GTAGATCTTAGCAGGGTGAAGGTAAAGTACTGTTACACGATAGATGTTGGAGCAACACAG 3062
Qy 646 -----PheGlnPro-----ValAlaAla 651
Db 3063 AGTGTAAAGTGAACAAGCAGCATAGATCCTCGGTATATAGATGTGAAGTTTGTGAAGCTT 3122
Qy 652 GlyLeuProSerSerGlyVala-----ValGly 660
Db 3123 GGAGCAACAGCAGCGGAGCGGATTAATCTATGTGAGATAGAGCTTTAAGAGTGGAGCAGCG 3182
Qy 661 ValMetPheHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGly 680
Db 3183 GTTTTGGCAGCAGGGCAAGAACACGAGGAGATAGACTTATGATACAGAAAGGCGAGTGGC 3242

Qy 681 LeuTyrHisSerThrAsnGlyGlySerSerTrpSerAla----- 693
Db 3243 AGCTCAATCACTCAAAATGACTATTTCGGTGAAGAGTGCACACAGGCTATATAGAGAACGAG 3302
Qy 694 ---IleThrGlyValSerSerAlaValAsnValGlyPheGlyLysSerAlaProGlySer 712
Db 3303 AAGGTAAACGGGTATATAGATGATGATCTTGTATGGGGAAGAGCGGACGAGCAACGCC 3362
Qy 713 SerTyrProAlaValPheValValGlyThrIleGlyGlyValThrGlyAla----- 729
Db 3363 CAGATCAAGGTATGTTATCGGAATGGATTTAAGCAGCCGACGAATGTATTGAATCCT 3422
Qy 730 ---TyrArgSerAspAspCysGlyThr----- 738
Db 3423 AAGATAAAGATAGAGAATTTGGGACGACACGCGGTAGATCTTTAGCAGGGTGAAGGTAAGA 3482
Qy 739 ---TrpValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThr 757
Db 3483 TACTGGTACACATAGATGGTGAGGCAACACAGAGTGTAAAGTGTAAACAGCAGCATAAAC 3542
Qy 758 GlyAspHisAlaAsnLeuArgValTyrIleGlyThrAsnGlyArgGlyIle---Val 776
Db 3543 CCTGCGTATATAGATGTGAAGTTTGTGAAGCTTGGAGCAANTGCCAGTGGAGCGGATTAC 3602
Qy 777 TyrGlyAspIleGly-----GlyAla----- 783
Db 3603 TATGTGGAGATAGGCTTTTAAGAGTGGAGCAGGGGTTTTGGCAGCAGGGCAGCAGCAAG 3662
Qy 784 -----ProSerGlySer----- 787
Db 3663 GAGATAAGACTTTAGCATACAGAAGGCGAGTGCAGCTACAANTCAGTCAAAATGACTATTTCG 3722
Qy 787 ----- 787
Db 3723 GTGAGGAGTGCACAGGCTATATAGAGAACAGAGGTAAACGGGTATATAGATGTGCG 3782
Qy 788 -----ProSerProSerValSerProSerAlaSerProSerLeuSer 801
Db 3783 ATAGTGTGGGGAAGAGAGCCGAGAGGGGTACAAAGCCGCGGGAGTAGTAAACCCGACA 3842
Qy 802 ProSerProSerProSerSerProSerProSerProSerProSerProSerProSer 821
Db 3843 CCGGCACCGACCCCGACATCGACGCGCACCAACACACTTACAACCACTGACCGCAGCA 3902
Qy 822 SerSerProSerProSerProSerProSerProSerProSerProSerProSerProSer 841
Db 3903 TCAGCCCGACACCCGAGCCCAACAGTGCACAGCAACCGCCGACTCCAACGCGCAGCGACA 3962
Qy 842 AlaSerProSerProSerProSerProSerProSerProSerProSerProSerProSer 861
Db 3963 GTGAGGTTACTGTGACTCCGACACCGACCA-----ACACCGACG 4004
Qy 862 ProThrProSerSerSerProValSerGly---GlyValLysValGlnTyrLysAsnAsn 880
Db 4005 CCGACACCGACAGGACACCTGCGCAGGAGTGGTTGAGGTACTATACAGAACAAAT 4064
Qy 881 AspSerAlaProGlyAspAsnGlnIleLysProGlyLeuGlnValValAsnThrGlySer 900
Db 4065 GAGACAAGTGGCAGCACAAAGTCTATAGGCGGTGTTTAAAGATAGTAGTAATGGAGCGAGC 4124
Qy 901 SerSerValAspLeuSerThrValThrValArgTyrTrpPheThrArgAspGlyGlySer 920
Db 4125 AGCAGTGTGTATCTTAGCAGGGTTAAGATAAGATCTGTTACACATGATGATGTGTGCAAG 4184
Qy 921 SerThrLeuValTyrAsnCysAspTrpAlaIleGlyCysGlyAsnIleArgAlaSer 940
Db 4185 CCACAGAGTGGGTA---TGTCACCTGGGCACAGATAGGGGCAAGCAATGTGACATTCAAT 4241
Qy 941 PheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 956
Db 4242 TTTTGTGAAGCTGAGCAGCGGAGTGTAGTGGCGGGATTATTACTTTGGAG 4289

RESULT 2

US-09-533-559-7511
; Sequence 7511, Application US/09533559
; Patent No. 6902887
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjarke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 5849-200-US
; CURRENT APPLICATION NUMBER: US/09/533,559
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 09/273,623
; EARLIER FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7511
; LENGTH: 1103
; TYPE: DNA
; ORGANISM: Tricoderma reesei
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1103)
; OTHER INFORMATION: n = A, T, C or G
US-09-533-559-7511

Alignment Scores:

Pred. No.:	2,44e-11	Length:	1103
Score:	392.50	Matches:	114
Percent Similarity:	45.0%	Conservative:	36
Best Local Similarity:	34.2%	Mismatches:	125
Query Match:	7.6%	Indels:	58
DB:	3	Gaps:	10

US-09-917-376-1 (1-957) x US-09-533-559-7511 (1-1103)

QY 348 GlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTrp 367
Db 3 GCGCTGGCGCTCGAATTTGCAAAAGCCAGGAAACCTTTGTTGCTTCTTGAACCTCTTGG 62
QY 368 TrpProAspThrIlePheArgSerThrAspGlyGlyAlaThrTrpThrArgIleTrp 387
Db 63 TGGCCAGATGCTCAGCTGTTTCGTCGACGACTCTGGGCAACATGGAGCCCGATCTGG 122
QY 388 AspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGluPro 407
Db 123 GCGTGGCGGAGCTATCCGACTGAGACCTATTACTACGACATCTCAACTCCCAAGCACCG 182
QY 408 TrpLeuThrPheGly---ValGlnProAsnProProValProSer-----Pro 422
Db 183 TGGATCAAGAACAACATTTATGATGTGACGAGCGAGTCAACCGTCCGATGCTCTNATCAAG 242
QY 423 LysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeu 442
Db 243 CGCCTCGGCTGGATGATGATGCTTCGAGATTGACCCACCCACCAANNACTGGCTT 302
QY 443 -TyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSer----- 460
Db 303 TTACGGCAGCGGATGACATNTTTGGCGGCCACGATTTCCCACTGGGACACGCGCCC 362
QY 461 -----GlyGlyGlnIleHisIleAlaProMetValLys-GlyLeuGluGluThrAlaVal 478
Db 363 ACAATGTGTCAATCAATTAATTTGTCGACAGCGGATTCGAAGGAATT-----TTTCGGT 416
QY 478 alaAsnAspLeuIleSerProProSerGly----- 487
Db 417 TCAAGGACCTGGCGCTTTTTCACCCGGGGGAAGCGAGCTTTTGGCCGCAAGTCCGGAGACG 476
QY 488 -----AlaProLeuIleSerAlaLeuAspLeuGlyGlyPheThrHisAlaAspV 505

Db 477 ANCAACGGGTTTACCTTTTGCCGACGAGAAACGACCTTGGG----- 517
QY 505 alThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThr-SerVal 524
Db 518 -----ACATTGCGCGCAGACGGTTGGGCAACTCCCAATGCGCCACCTCGACGAAGCGTC 572
QY 525 AspTyrAla--GluLeuAsnProSerIleIleValArgAlaGlySer----- 539
Db 573 GACTACGCGCGGAACTCGTCAAGAGCGGTTCTCGTCCGCGTGGCAACACCGCGCGGCACG 632
QY 540 -----PheAspProSerSerGlnProAsnAspArgHisValAlaPheS 554
Db 633 CAACAAGGTGGCCATTTTTCGAAACGCGCGCGCGACGATCTCTGGTCACCGCTCGTCGGC- 686
QY 554 erThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGlyValThrThrGlyG 574
Db 687 GAACATACGCT-----GGNTCCGAACACCGTTTCCATTGAAACGGCG 728
QY 574 lyThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaProGlyAspProGlyG 594
Db 729 GCGCGTGGCTATTTCGCGCAGCGCGACGATCTCTGGTCACCGCTCGTCGGC- 787
QY 594 lnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProA 614
Db 788 -----GTGACGCGCTCGCAGTTTCCAGGGCAGCTTTGCTCTCGAGCCTCGCG 839
QY 614 laAsnAlaGlnIleArgSerAspArgValAsnProLysThrPheTyrAlaLeuSerAsnG 634
Db 840 CGGGCGCGCTCATCGNCTCGGACAAAGAGACCAACAGCGTNTTCTACGCGCGCTCGGAT 899
QY 634 lyThrPheTyrArgSerThrAspGlyGly 643
Db 900 CGACCTTTTACGTACGACGACACCGCGC 928

RESULT 3

US-09-949-016-16031/c
; Sequence 16031, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16031
; LENGTH: 19383
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(19383)
; OTHER INFORMATION: n = A, T, C or G
US-09-949-016-16031

Alignment Scores:

Pred. No.:	3.36e-08	Length:	19383
Score:	358.00	Matches:	71
Percent Similarity:	84.9%	Conservative:	2
Best Local Similarity:	82.6%	Mismatches:	13
Query Match:	7.0%	Indels:	0
DB:	3	Gaps:	0

US-09-917-376-1 (1-957) x US-09-949-016-16031 (1-19383)


```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 18597
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 701..13751
; OTHER INFORMATION: n = c or a
; NAME/KEY: misc_feature
; LOCATION: 716..1293, 2401..2429, 2618, 3083, 3125, 3635, 4256, 4898,
; LOCATION: 5062, 5167, 11069, 13298, 14479, 14730, 14796, 15344, 15450,
; LOCATION: 15503, 15590, 15840, 16149
; OTHER INFORMATION: n = a or g
; NAME/KEY: misc_feature
; LOCATION: 732, 1379, 1590, 2488, 3212, 5006, 11238, 11422, 11686,
; LOCATION: 12598, 13171, 13645, 13782, 13806, 13813, 14586, 14788,
; LOCATION: 15042, 15546, 15770
; OTHER INFORMATION: n = c or t
; NAME/KEY: misc_feature
; LOCATION: 1322, 1688
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 2594..11293, 16199, 16203
; OTHER INFORMATION: n = g or t
; NAME/KEY: misc_feature
; LOCATION: 3619
; OTHER INFORMATION: n = a or t
; NAME/KEY: misc_feature
; LOCATION: 14547
; OTHER INFORMATION: nucleotide in position 14547 is t, or absent
US-09-962-665-8

Alignment Scores:
Pred. No.:      1.42e-07      Length:      18597
Score:          345.00        Matches:     71
Percent Similarity: 83.9%    Conservative: 2
Best Local Similarity: 81.6% Mismatches:   13
Query Match:      6.7%       Indels:      1
DB:               3         Gaps:        0

US-09-917-376-1 (1-957) x US-09-962-665-8 (1-18597)
Qy 783 AlaProSerGlySerProSerProSerValSerProSerAlaSerProSerIeuSerPro 802
:::|||||
Db 10388 TCACCATCACCATTCTCCATCTCCATCACCATCACCATCACCATCACCATCACCA 10359
Qy 803 SerProSerProSerSerSerProSerProSerProSerProSerSerSerProSerSer 822
|||:|||||
Db 10328 TCACCATCTCCATCCACCATCACCATCACCATCTCCATCACCATCACCATCTCCATCACCA 10269
Qy 823 SerProSerProSerProSerProSerProSerProSerProSerArgSerProSerAla 842
|||:|||||
Db 10268 TCACCATCACCATCACCATCTTCATCACCATCACCATCACCATCTCCATCACCATCACCA 10209
Qy 843 SerProSerProSer-SerSerProSerProSerProSerSerProSerSerProSerPr 862
|||:|||||
Db 10208 TCACCATCACCATCCCCATCACCATCACCATCACCATCACCATCTCCATCACCATCACCA 10149
Qy 862 OThrProSerSerSerPro 868
|||:|||||
Db 10148 ATCTCCATCACCATCACCA 10130

RESULT 7
US-09-963-333-8/c
; Sequence 8, Application US/09963333
; Patent No. 6664062
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES
; TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT
; TITLE OF INVENTION: OF DISEASE
; FILE REFERENCE: 11926-015002
```

```

CURRENT APPLICATION NUMBER: US-09/963,333
CURRENT FILING DATE: 2001-09-24
PRIOR APPLICATION NUMBER: 09/658,659
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 09/596,033
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 09/357,743
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 09/357,024
PRIOR FILING DATE: 1999-07-19
PRIOR APPLICATION NUMBER: 60/093,484
PRIOR FILING DATE: 1998-07-20
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 18597
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 701..13751
OTHER INFORMATION: n = c or a
NAME/KEY: misc feature
LOCATION: 716..1293, 2401..2429, 2618..3083, 3125..3635, 4256..4898,
LOCATION: 5062..5167, 11069..13298, 14479..14730, 14796..15344, 15450,
LOCATION: 15503..15590, 15840..16149
OTHER INFORMATION: n = a or g
NAME/KEY: misc feature
LOCATION: 732..1379, 1590..2488, 3212..5006, 11238..11422, 11686,
LOCATION: 12598..13171, 13645..13782, 13806..13813, 14586..14788,
LOCATION: 15042..15546, 15770
OTHER INFORMATION: n = c or t
NAME/KEY: misc feature
LOCATION: 1322..1688
OTHER INFORMATION: n = c or g
NAME/KEY: misc feature
LOCATION: 2594..11293, 16199..16203
OTHER INFORMATION: n = g or t
NAME/KEY: misc feature
LOCATION: 3619..
OTHER INFORMATION: n = a or t
NAME/KEY: misc feature
LOCATION: 14547
OTHER INFORMATION: nucleotide in position 14547 is t, or absent
US-09-963-333-8

Alignment Scores:
Pred. No.: 1.42e-07 Length: 18597
Score: 345.00 Matches: 71
Percent Similarity: 83.9% Conservative: 2
Best Local Similarity: 81.6% Mismatches: 13
Query Match: 6.7% Indels: 1
DB: 3 Gaps: 0

US-09-917-376-1 (1-957) x US-09-963-333-8 (1-18597)
Qy 783 AlaProSerGlySerProSerProSerValSerProSerAlaSerProSerLeuSerPro 802
Db TCACCATCACCATCTCCATCTCCATCACCATCACCATCACCATCACCATCACCA 10329
Qy 803 SerProSerProSerSerProSerProSerProSerProSerProSerSerSer 822
Db TCACCATCTCCATCACCATCACCATCACCATCTCCATCACCATCACCATCACCA 10269
Qy 823 SerProSerProSerProSerProSerProSerProSerArgSerProSerAla 842
Db TCACCATCACCATCACCATCTCCATCACCATCACCATCACCATCTCCATCACCA 10209
Qy 843 SerProSerProSer-SerSerProSerProSerSerProSerSerProSerPr 862
Db TCACCATCACCATCACCATCACCATCACCATCACCATCACCATCTCCATCACCA 10149
Qy 862 oThrProSerSerPro 868

```

Db 10148 ATCTCCATCACCATCACCA 10130
|:::|||||

RESULT 8

US-09-962-677-8/c
; Sequence 8, Application US/09962677
; Patent No. 6759200
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE PHOSPHORYLASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING
; TITLE OF INVENTION: THE TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015003
; CURRENT APPLICATION NUMBER: US/09/962.677
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,459
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 18597
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 701..13751
; OTHER INFORMATION: n = c or a
; NAME/KEY: misc feature
; LOCATION: 716, 1293, 2401, 2429, 2618, 3083, 3125, 3635, 4256, 4898,
; LOCATION: 5062, 5167, 11069, 13298, 14479, 14730, 14796, 15344, 15450,
; LOCATION: 15503, 15590, 15840, 16149
; OTHER INFORMATION: n = a or g
; NAME/KEY: misc feature
; LOCATION: 732, 1379, 1590, 2488, 3212, 5006, 11238, 11422, 11686,
; LOCATION: 12598, 13171, 13645, 13782, 13806, 13813, 14586, 14788,
; LOCATION: 15042, 15546, 15770
; OTHER INFORMATION: n = c or t
; NAME/KEY: misc feature
; LOCATION: 1322, 1688
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc feature
; LOCATION: 2594, 11293, 16199, 16203
; OTHER INFORMATION: n = g or t
; NAME/KEY: misc feature
; LOCATION: 3619
; OTHER INFORMATION: n = a or t
; NAME/KEY: misc feature
; LOCATION: 14547
; OTHER INFORMATION: nucleotide in position 14547 is t, or absent

US-09-962-677-8

Alignment Scores:

Pred. No.: 1.42e-07 Length: 18597
Score: 345.00 Matches: 71
Percent Similarity: 83.9% Conservatives: 2
Best Local Similarity: 81.6% Mismatches: 13
Query Match: 6.7% Indels: 1
DB: 3 Gaps: 0

US-09-917-376-1 (1-957) x US-09-962-677-8 (1-18597)

Qy 783 AlaProSerGlySerProSerValSerProSerAlaSerProSerLeuSerPro 802
|:::|||||

Db 10388 TCACCATCACCATCTCCATCTCCATCACCATCACCATCACCATCACCATCACCA 10329

Qy 803 SerProSerProSerSerProSerProSerProSerProSerProSerProSer 822
|:::|||||

Db 10328 TCACCATCTCCATCACCATCACCATCACCATCTCCATCACCATCTCCATCACCA 10269

Qy 823 SerProSerProSerProSerProSerProSerProSerProSerProSerAla 842
|:::|||||

Db 10268 TCACCATCACCATCACCATCTCCATCACCATCACCATCTCCATCACCATCACCA 10209

Qy 843 SerProSerProSerProSerProSerProSerProSerProSerProSerProSerPr 862
|:::|||||

Db 10208 TCACCATCACCATCACCATCACCATCACCATCACCATCACCATCTCCATCACCAC 10149

Qy 862 oThrProSerSerPro 868
|:::|||||

Db 10148 ATCTCCATCACCATCACCA 10130

RESULT 9

US-09-136-574A-2

; Sequence 2, Application US/09136574A

; Patent No. 6294366

; GENERAL INFORMATION:

; APPLICANT: Farrington, Graham K.

; Anderson, Paige

; Gibbs, Moreland

; Bergquist, Peter

; Daniels, Roy

; Morgan, Hugh W.

; Williams, Diane P.

; TITLE OF INVENTION: Compositions and Methods for

; Treating Cellulose Containing Fabrics Using Truncated

; Cellulase Enzyme Compositions

; NUMBER OF SEQUENCES: 49

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Howson and Howson

; STREET: Spring House Corporate Center, P.O. Box 457

; CITY: Spring House

; STATE: PA

; COUNTRY: USA

; ZIP: 19477

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/136,574A

; FILING DATE: 19-Aug-1998

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/932,571

; FILING DATE: September 19, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Bak, Mary E.

; REGISTRATION NUMBER: 31,215

; REFERENCE/DOCKET NUMBER: 1997US001/CIP

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 215-540-9200

; TELEFAX: 215-540-5818

; TELEX: <Unknown>

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 6416 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-136-574A-2

Alignment Scores:

Pred. No.: 1.49e-07 Length: 6416
Score: 334.00 Matches: 219
Percent Similarity: 33.3% Conservatives: 125
Best Local Similarity: 21.2% Mismatches: 314
Query Match: 6.5% Indels: 376

DB: 3 Gaps: 56
 US-09-917-376-1 (1-957) x US-09-136-574A-2 (1-6416)
 QY 18 SerLeuLeuAlaAlaThrAlaSerPheAlaValAlaAlaLeuGlyValLeuProIle 37
 Db 1189 TCTGTAGTGCAGAACCA-----GCTGCATCAGTTCGGGTTCAGTT 1233
 QY 38 AlaIleThrAlaSerProAlaHisAlaAla----- 47
 Db 1234 GTAATAAGAGAAAGAAATCTCAGAAAGCAGCTTCTTATCTCCAAATGTCACCAAGACCTG 1293
 QY 48 -----ThrThrGlnPro-----TyrThrTrpSerAsnValAlaIle 59
 Db 1294 TTTGAATTTGCCGATACCAAGAAGTGTGCGGGTATATCTGTCGCAACAGGTTTCTAC 1353
 QY 60 GlyGlyGlyGlyPheValaspGlyIleValPheAsnGluGlyAlaProGlyIleLeuTyr 79
 Db 1354 ACATCGGTGGTTTTATGTATGACCTTGGATGG-----GCTGCTGTATGGCTTTAT 1404
 QY 80 ValArgThrAsp----- 84
 Db 1405 ATTGCGCAAAATCAGACAGTATTATTCACGAAGCTGAAGAGTTGATGTCAGATATGCT 1464
 QY 85 GlyGlyMetTyrArgTrp-----AspAlaIleAsnGlyArgTrpIlePro 99
 Db 1465 AATGGTACTAATACATGACGACCAATGCTGGGATGATGTCGATATGCAACATTGATCATG 1524
 QY 100 LeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAla 119
 Db 1525 CTTGCAAGATTACAGGAAAGATT-----TATAAGGAGCTGTGGAAAGAACTTA 1578
 QY 120 AspProIleAsnThrAsnLysValTrpAlaAlaValGlyMet---TyrThrAsnSerTrp 138
 Db 1579 GACATTGGACTACAGAAATTAGTATACCGCAAGGGATGGCATATCTGACAGGATGG 1638
 QY 139 AspProAsnAspGlyAlaIleLeuArgSerSerAspGlnGly----- 152
 Db 1639 -----GGTTCAATTAAGATATGCGCAACACAGCTGCATTTTATGATGTGCTAT 1686
 QY 153 AlaThrTrpGln----- 158
 Db 1687 GCAGACTGTGTCAGGGTGCAGTTCGAAACAAAGACCAAAATATTGAACTTTGCAAAAGC 1746
 QY 159 ProLeuProPheLysLeuGlyGlyAsnMetProGlyArg-----GlyMetGly 174
 Db 1747 CAGATTGACTGCACCTGGGTCCACA-----GGTAGAGTTTGTAGTAGGATTTGGC 1800
 QY 175 GluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGly 194
 Db 1801 ACCAATTATCCACACATCGCATCAGCAAT----- 1833
 QY 195 LysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPhePro 214
 Db 1834 -----GGCAGATGTCATGGCTAACGATGATAAAATACCA 1869
 QY 215 Asp-----ValGlyThrTyrIleAlaAsnProThrAspThrThrGly 228
 Db 1870 GAGTATACAGACACATATTATATGAGCAGCTGTTGGTGGTCTCTGTAGTATGATAGT 1929
 QY 229 TyrGlnSerAspIleGlnGlyValValTrp-----ValAlaPheAspLysSerSer 246
 Db 1930 TATAATCATGACATTACCGATTATGACAAATGAGTTTCCCTGCGCATTAATATCTGGA 1989
 QY 247 SerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspProAsnAsnProVal 266
 Db 1990 ATTGTTGGTGCACTGGCAAG----- 2010
 QY 267 PheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPhe 286
 Db 2011 -----ATGTACCAGTTATATGAGGTGAACCTATTGAT--- 2043
 QY 287 IleProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThr----- 304

Db 2044 ---GATTTTAAAGCAATTTGAAACACCCCAAAATGATGAAATTTTGTTCATCAAAATTT 2100
 QY 305 SerAsnThrGlyGlyProTyrAspGlySerSerGlyAspValTrpLysPheSerValThr 324
 Db 2101 GGGAAATTCAGAGGTCCA-----AATTATACCGAAGTAATTTCTTATATCTATAAT 2151
 QY 325 SerGlyThrTrpThrArgIleSerProValProSerThrAspThrAlaAsn---AspTyr 343
 Db 2152 CGAACAGATGS-----CCACCAAGGGTAACCTGATAAACTAAGTTTAAATAT 2199
 QY 344 Phe-----GlyTyrSerGlyLeuThrIleAspArgGln 354
 Db 2200 TTTATAGACCTTAACCGAATTAATCCAGGCAGGGTATTCG----- 2238
 QY 355 HisProAsnThrIleMetValAlaThrGlnIleSerTrpTrpProAspThrIleIlePhe 374
 Db 2239 ---CCTGATGTTGTCAAGTTTGACACA-----TAC 2265
 QY 375 ArgSerThrAspGlyGlyAlaThrTrpThrArgIleTrpAspTrpThrSerTyrProAsn 394
 Db 2266 TACATCGAAGGAGGTAAATTTAGCGGTCTTACGTATGGAC-----AAAAAT 2313
 QY 395 ArgSerLeuArgTyrValLeu---AspIleSerAlaGluProTrpLeuThrPheGlyVal 413
 Db 2314 AGGAATATATATATGTTCTTGTGATTTTGTAGTGAAACCAAG-----ATA 2358
 QY 414 GlnProAsnProProValProSerProLysLeuGlyTrp-----MetAspGlu 429
 Db 2359 TATCCTGGCGGTGAAGTTGAACACAAAGACGCGCTCAATTTAAATATCTGTCGCGAG 2418
 QY 430 AlaMetAlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThrGlyAlaThrLeu 449
 Db 2419 GGGTATCCATGGATCTTACCAAT---GATCCTCATATAGGATTAACAGTCAATTA 2475
 QY 450 TyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlyGlnIleHisIleAlaProMet 469
 Db 2476 GAAAGAATAAATAATATTGCGCATATGATAATAATAAT-----CTG 2517
 QY 470 ValLysGlyLeuGluGluThrAlaValAsnAspLeuIleSerPro----- 484
 Db 2518 GTATGGGGTTTTAGAGCGGGTGCAGCAACATCCACCTGACCAACATCAACACCAACA 2577
 QY 485 -----ProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHis 502
 Db 2578 CCAACCCCGACCCCAACACCAACAGTGCAGCA----- 2610
 QY 503 AlaAspValThrAlaValProSerThrIlePheThr-----SerProValPheThr 519
 Db 2611 -----ACGCCGACCGCGACTCTTACACCGACACCGCGGGTCACT---GGT 2655
 QY 520 ThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySer 539
 Db 2656 ACGGGAAGTGTGTGAAGGTACTGTACAGAACAATGAGCAAGTGCAGCAGCAGGTCT 2715
 QY 540 PheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyLys 559
 Db 2716 ATAAGGCCG----- 2724
 QY 560 AsnTrpPheGlnGlySerGluProGlyValThrThrGlyGlyThrValAlaAlaSer 579
 Db 2725 ---TGGTTTAAG-----ATAGTGAATGGAGGC-----AGCAGCAGT 2757
 QY 580 AlaAspGlySerArgPhe-----ValTrp-----AlaProGlyAspProGlyGln 594
 Db 2758 GTTGATCTTAGCAGGGTTAAGATAAGTACTGGTACAGTGGTGGTGACAGCCACAG 2817
 QY 595 ProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAla 614
 Db 2818 AGTGGCGTATGT-----CACTGG----- 2835
 QY 615 AsnAlaGlnIleArgSerAspArgValAsnProLysThrPheTyrAlaLeuSerAsnGly 634

```

Db 2836 ---GCACAGATAGGGCGCAAGCAATCTGACA---TTCAATTTTGTGAAGCTTAGCAGCGGA 2889
Qy 635 ThrPheTyrArgSerThrAspGlyValThrPheGlnProValAlaAlaGlyLeuPro 654
Db 2890 GTG-----AGTGGAGCGGATTATTAC---CTGGAGGTAGGATTT---2925
Qy 655 SerSerGlyAlaValGlyValMetPheHisAlaValProGlyLysGlu---GlyAspLeu 673
Db 2926 ACAGTGGAGCTGGCGAGTTGCGAG-----CTGTGTAAGGACACAGGGGATATA 2973
Qy 674 TrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsnGlyGlySerSerTrpSerAla 693
Db 2974 CAGGTAAGGTTTAAACAAGATCACTGGAGCAATACATCAGGCAGACACTGGTTCATGG 3033
Qy 694 IleThrGlyValSerSerAlaValAsnValGlyPheGlyLysSerAlaProGlySerSer 713
Db 3034 TTGCAGAGCATGACCAAT-----TATGAGAGAAATGCCAAGGTGACGCTG 3078
Qy 714 TyrProAlaValPheValValGlyThrIleGlyGlyValThrGlyAlaTyrArgSerAsp 733
Db 3079 TAT-----GTAGATGGT-----3090
Qy 734 AspCysGlyThrThrTrpValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnTrpGly 753
Db 3091 -----GTCTGTGTA-----TGGGGG 3105
Qy 754 GlnAlaIleThrGlyAspHisAlaAsnLeuArgValTyrIleGlyThrAsnGlyArg 773
Db 3106 CAG-----3108
Qy 774 GlyIleValTyrGlyAspIleGlyAlaProSerGlySerProSerProSerValSer 793
Db 3109 -----GAGCGGAGGAGCG-----3123
Qy 794 ProSerAlaSerProSerLeuSerProSerProSerProSerSerProSerProSer 813
Db 3124 -----ACACTGCACCGCAACGACACACCAACCAACG 3156
Qy 814 ProSerProSerSerProSerSerProSerProSerProSerProSerProSerProSer 833
Db 3157 CCAACTCCGACGACGACCAACCCACCTACCTACCGCCGACCCCGACACCGACAGTGAGT 3216
Qy 834 ProSerArgSerProSerProSerSerProSerProSerProSerProSerProSerProSer 853
Db 3217 GCAACGCCAACACACCGCACCGGCATCACCGTAGTGGCAGTTACTGGACGCCGAGT 3276
Qy 854 SerSerProSerSerProSerProSerProSerProSerProSerProSerProSerGlyVal 873
Db 3277 GAGAGT-----TACGGTGGCTG 3294
Qy 874 LysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnIleLysProGlyLeu 893
Db 3295 AAGGTATGTTATGCGAATGGGAATTTAGACGCCGACGAATGATTAATCCTAAGATA 3354
Qy 894 GlnValAlaAsnThrGlySerSerSerValAspLeuSerThrValThrValArgTyrTrp 913
Db 3355 AGATAGAGATGTTGGGACGACGACGCGTAGATCTTACAGGTTGAAGTAACTAGTGG 3414
Qy 914 PheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaIleGly 933
Db 3415 TACACGATAGATGTTGAGCGCACACAGAGTGTA-----3447
Qy 934 CysGlyAsnIleArgAlaSerPheGlySerValAsnProAla 947
Db 3448 -----AGTGTAGCGAGCAGCATAAATCTCTGCG 3474

```

RESULT 10

```

US-09-252-991A-13656
; Sequence 13656, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

```

```

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

```

```

; FILE REFERENCE: 107196.136

```

```

; CURRENT APPLICATION NUMBER: US/09/252,991A

```

```

; CURRENT FILING DATE: 1999-02-18

```

```

; PRIOR APPLICATION NUMBER: US 60/074,788

```

```

; PRIOR FILING DATE: 1998-02-18

```

```

; PRIOR APPLICATION NUMBER: US 60/094,190

```

```

; PRIOR FILING DATE: 1998-07-27

```

```

; NUMBER OF SEQ ID NOS: 33142

```

```

; SEQ ID NO 13656

```

```

; LENGTH: 8211

```

```

; TYPE: DNA

```

```

; ORGANISM: Pseudomonas aeruginosa

```

```

US-09-252-991A-13656

```

Alignment Scores:

```

Pred. No.: 1,046-06 Length: 8211
Score: 319.50 Matches: 290
Percent Similarity: 32.6% Conservatives: 133
Best Local Similarity: 22.3% Mismatches: 449
Query Match: 6.2% Indels: 432
DB: 3 Gaps: 69

```

```

US-09-917-376-1 (1-957) x US-09-252-991A-13656 (1-8211)

```

```

Qy 18 SerLeuLeuAlaAlaThrAlaSerPheAlaValAlaAlaLeuGlyValLeuProIle 37
Db 3326 GCAATGGCAGCAGCCTCAGCGGCATCGCGGAACCGGCGAGCAGCGGTGATCTCACCACG 3385
Qy 38 AlaIleThr-----AlaSerProAlaHisAlaAlaThrThrGlnPro-----51
Db 3386 GCAACGGCAATCCGATCGCGGAGTTCACCGCGCAGCGCAGCGCAACTGCACCTACACCC 3445
Qy 52 -----TyrThrTrpSerAsnVal-----AlaIleGly 60
Db 3446 GTCCACGCCGATCCCAACGGTACTGTGTC--AACGTGGTGGCCGAGACGCCCGCGGT 3504
Qy 61 GlyGlyGlyPheValAspGlyIleValPheAsnGluGlyAlaProGlyIleLeuTyrVal 80
Db 3505 AACACGAGCCCGCGGAGCGGTGACCGTCGATTCAGCGCGCGCGCGCGCGGTGATC 3564
Qy 81 ArgThrAspIleGlyMetTyrArgTrpAspAlaAlaAsnGlyArgTrpIleProLeu 100
Db 3565 AACCCGAGCAACGGCGTGTCTATCAGCGCACCGCGGAGCGCGGTGCCACCGTGACCTC 3624
Qy 101 LeuAspTrpValGlyTrpAsn-----AsnTrp 109
Db 3625 ACCGAT---GCCGCGGCAACCCGATAGGGCAGGTACACCGCGACCGCGCAACTGG 3681
Qy 110 GlyTyr-----AsnGly-----ValValSerIleAlaAlaAsp 120
Db 3682 AGCTTCACGCCGGGACGCCCGCGGCAACGGCAGCGGTGATGTGCCACGCCACCGAC 3741
Qy 121 Profile-----AsnThrAsnLysValTrpAla-----AlaValGlyMetTyrThr 135
Db 3742 CCGACCGGCAATACCGCGCGGCGGCGGCGCACACCGGTGACGCGTGGCGCGCGCGCG 3801
Qy 136 AsnSerTrpAspProAsnAspGlyAlaIleLeu-----146
Db 3802 CCGGTGATCGATCCGAGCAACCGCACGACCATCAGCGGCACCGCGGAGCGCGGCGCAAG 3861
Qy 147 -----ArgSerSerAspGlnGlyAla 153
Db 3862 GTGATCTTCACCGAGCGCAACCGCAACCGATCGCGGAAACCAACCCCGCGACGCGGC 3921
Qy 154 ThrTrpGlnIleThrPro-----LeuProPheLysLeuGly-----165
Db 3922 AACTGACCTTCACCGCGCGCGCGCGTGGCGCAACGGCACCGGTGTCAACGCCGTGGCC 3981
Qy 166 -----GlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspPro 181
Db 3982 CAGGACCTTGGCGGCAATACCGCGCGCGGAGCGAGCAGCACTACCGTGGACGCGGTGGCGCG 4041

```

Qy 182 Asn-----AsnAspAsnIleLeuTyrPheGlyAlaProSerGly 194
Db 4042 AACAGCGCTGGTCAATCCGAGCAACGGCAACCTGCTCAACGGTACCGCGGAGCCGGC 4101
Qy 195 LysGlyLeu-----TrpArgSerThrAspSer 203
Db 4102 AGCACCCTGACCTTGACCGAGCGCAACGGCAACCCGATCGGCCAGACACCCCGCATGGC 4161
Qy 204 GlyAlaThrTrp-----SerGlnMetThrAsnPheProAspValGlyThr 218
Db 4162 AGCGGCAACTGGAGCTTCACGCCCGGCTCGCACTACCCAC-----GGCACC 4209
Qy 219 TyrIle---AlaAsnProThrAspThrThrGlyTyrGlnSerAspIleGlnGlyValVal 237
Db 4210 GTGGTCAACGTGACCGGAGCGCGCGGCAATACACG-----GCTCCCGCT 4260
Qy 238 TrpValAlaPheAspLysSerSerSerLeuGlyGln-----AlaSerLysThr 254
Db 4261 ACCAGAGGTGGATTCCTCGCTCCGTTCGATCCCGAGGTGGATCCGAGCAACGGTTCG 4320
Qy 255 IlePheValGlyValAlaAspProAsnProValPhe----- 267
Db 4321 GTGATCAGCGCACCGCGGCGCGCAACACCATCATCATCACCAGTGGCAACGGCAAC 4380
Qy 268 -----TrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaPro 283
Db 4381 CCGATTGGCCAGGTACCGCGCGAGCGGAGCGGTAACCTGCTTCCTTCACTCCAGGCATCCG 4440
Qy 283 ----- 283
Db 4441 CTGCCGATGGCACGGTGTCAACGTGGTGGCGCGCACGCCCAAGCAATGTCAGATGGC 4500
Qy 284 -----ThrGlyPheIleProHisLysGlyValPheAspProValAsn 297
Db 4501 CCGCGGTGATCACTGTGGATGGCGTGGCCCGCGCGCGCGGTGATCGATCCAGCAAC 4560
Qy 298 HisValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyr-----AspGlySer 314
Db 4561 GGCACCGAGATAAGCGGTACCGCGGAGCGCGCGCGGTGATCTCTCACCGATGGCGGC 4620
Qy 315 SerGlyAspValTrpLysPheSerValThr---SerGlyThrTrpTrpArgIleSerPro 333
Db 4621 GGC AACCCGATCGGCCAGCGCACCGCGCGGAGCGGCACTGGAGCTTCACCCCGCGGC 4680
Qy 334 ValPro-----SerThrAspThrAlaAsnAspTyr 343
Db 4681 ACCCGCTGGCAACGGCACCGTGTATCAACCGCTGGCGCGAGGACCCCGCGCAATACC 4740
Qy 344 PheGlyTyrSerGlyLeuThrIleAspArgGlnHisPro----- 356
Db 4741 AGCGTCCGCGCAGCGTCACGTCGATGCCATCGCCCGCGCGCGGTGATCAATCCG 4800
Qy 357 ---AsnThrIleMetValAlaThrGlnIleSerTrpTrpProAspThrIleIlePheArg 375
Db 4801 AGCAACCGAGTGTGTCATCAGCGGTACCGGGAAGCGCGGCGCACCGTGTATCTC----- 4854
Qy 376 SerThrAspGlyGly-----AlaThrTrpThrArgIleTrpAspTrp 389
Db 4855 ---ACCAGCGCAACCGGCAACCCGATCGGCCAGGTTCACCGCGGAGCGCGCAAGTGG 4911
Qy 390 Thr-SerTyrProAsnArg-----SerLeuArgTyrValLeuAs 402
Db 4912 GCTTTCAGCCCGCACCGCGTTGGCCATGGCAGGTGATCATATGCGTGG----- 4963
Qy 402 pIleSerAlaGluProTrpLeuThrPheGlyValGlnProAsnProProValProSerPr 422
Db 4964 ---CCAGGACGCGCGCGCAACACAGCAGTC-----CCACCGCGCCA---CC 5007
Qy 422 OlysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLe 442
Db 5008 GTCACTCGTGGCGCCAGCA-GCCCGCGTGTATCGATCCGAGCAACCGGTAGC---GTGAT 5063
Qy 442 uTyrGlyThr-----GlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAs 459

Db 5064 CCCTGGTACCCCGGAGGCTGGTCCACGGTATCTCTACC-----GA 5105
Qy 459 pSerGlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluThrAlaValAs 479
Db 5106 CGGCAACGGCAAC-----CCGATCGGCAGGTACCCCGCATGGCAGCGGCA 5153
Qy 479 nAspLeuIleSerProSerGlyAlaProLeu-----IleSerAlaLe 494
Db 5154 CTGGAGCTTCACGCC-----GGCAGCGCTGTCCAATGGCACCGTGGTCAATCGGT 5207
Qy 494 uGly---AspLeuGlyGlyPheThrHisAlaAspVal---ThrAlaValProSerThrI 512
Db 5208 GGCCAGGACGCTGCGGCAACACACGAGCGCGCGCGCACACCGATCGGTGATCGGTGGC 5267
Qy 512 ePheThrSerProValPhe-----ThrThrGlyThrSerVal---AspTyrAlaGluLe 529
Db 5268 CCGCGCGCGCGGTGATCGACCGGCAACGGCAGCGTGTATCGCGGTACCGCGGAGC 5327
Qy 529 uAsnProSerIleValArgAlaGlySerPheAspProSerSerGlnProAsnAspAr 549
Db 5328 CGGTGCGAGGTGATCTCTCACCGATGGCGGCAACCGATCGGCCAGGCC----- 5379
Qy 549 gHisValAlaPheSerThrAspGlyGlyLysAsnTrp----- 561
Db 5380 -----ACCGCGATGGCAGCGCAACTGGAGCTTCACCCCGGCGCACCGCGT 5426
Qy 562 -----PheGlnGlySerLeuProGlyGlyValThrThrGlyG 574
Db 5427 GGCACACGGCACGGTGTATCAATCGGTGGCGCGCAGGATCCGCGCGCAATACCGCGGCC 5486
Qy 574 yThr----- 575
Db 5487 GACCAGCACACCGTGGAGCGGTGGCCCGCGCACCCCGGTGGTCAACCCGAGCAACGG 5546
Qy 576 -----ValAlaLaserAlaAspGlySerArgPheValTrpAlaProGlyAspProG 593
Db 5547 CAGCGTGTATCGCTACCGCGGAAGCGCGCGCGCGGTGATCTCTCACCGAGCGCGCGG 5606
Qy 593 yGlnProVal-----ValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGlnG 611
Db 5607 CAACCCGATCGCGCAGGTTCACCGCGCGAGCGGCGGCAAC---TGGAGCTTCACGCGCG 5663
Qy 611 yValPro---AlaAsnAlaGlnIleArgSer----- 620
Db 5664 CACCGCGTGGCCCAACCGCTCGGTGATCATCGGTGGCCAGGACGCGCGCGCAACAC 5723
Qy 621 -----AspArgValAsnProLysThr---PheTyrAlaLe 631
Db 5724 CAGCGCGCGCGCACCGACCGTGGTGTCTCGTAGCCCGCGCACCCCGGTGTCTGATCC 5783
Qy 631 uSerAsnGlyThrPheTyrArgSer-----ThrAs 641
Db 5784 GAGCAACCGTACCGGTGATCAGCGGTACCGCGCAAGCGCGGCGCACCGTGTCTCTCACCG 5843
Qy 641 pGlyGlyValThrPheGlnProValAlaAlaGlyLeu-ProSerSerGlyAlaValGlyV 661
Db 5844 CGGCGGC-----GGCAACCGGATACGGCAGCGCACCGCGCGCATGGCAGCGGCAACTGGAG 5897
Qy 661 alMetPheHisAlaValProGlyLysGlyAspLeuTrpLeuAlaAlaSerSerGlyL 681
Db 5998 CTTCACTCCGGGCACA-CCGCTGACCAACGCGACCGGTGATCAATCGGTGGGCCAGGACG 5956
Qy 681 euTyrHisSerThrAsnGly-----GlySerSerTrpSerAlaIleThrGlyValSerS 699
Db 5957 CCGCGCGCAACACCGAGCGGTCCGTCAGCACACAGTGGACGCGGTGGCGCGCGCACCC 6016
Qy 599 erAlaValAsnValGlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheV 719
Db 6017 CGGTGATCGACCGGCAATGTGTCAAACTCAGCGGCAACCGCGCGCGCGGTCTCCGGG 6076
Qy 719 alValGlyThrIleGlyGly----- 725

```

Db 6077 TGATCTCACCAGTGGCAATGCAACCGATCGGCCAGACCTCGCCGAGGTAGCGCA 6136
Qy 726 -----valThrGlyAlaTyrArgS 732
Db 6137 ACTGACCTTCACACCGGCGACCGCTGGCCAAACGCGACGCTGTCAACGCGGTGGCC 6196
Qy 732 eRAspAspCysGlyThrThrTrpValLeuLeuAsnAspGlnHisGlnTyrGlyAsnT 752
Db 6197 AGGACCGCGCGCNAATACAG-----CGTCCGCGCAGCAGCAGCAGGTGTGATACG 6246
Qy 752 rpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgArg-ValTyrIleGlyThrAsn 771
Db 6247 GTGGCGCGCGCACCGCGTGTATCAATCCAGCAACGCGAGCGTGTATCACCGCACCGCC 6306
Qy 772 -----GlyArgGlyIleValTyrGlyAspIleGlyAlaProSerGlySer-ProSe 789
Db 6307 GAGGTGGCGCGCAAAAGTGTCTCACCGCAGCGCAACCGCATCGCGGAGACCAACC 6366
Qy 789 rProSerVal-----SerProSerAlaSer-----ProSerLeu----- 800
Db 6367 GCCGAGCGCACTGGCAACTGACCTTCACCCCGCGCAGCGCGCTGGCCCAACGTCAGGTG 6426
Qy 801 -----SerProSerPro-----SerProSerSerProSerProSerPro----- 816
Db 6427 ATCAACCGCGTCCGCGAAGACCGCGCGGCAACGCGCGTCCGCGCAGCACCAACGCGT 6486
Qy 817 -----SerSerSerProSerSerSerProSerPr 826
Db 6487 GACTCGGTGGCGCGTCCGCTCCGCTGCTGAGCATCAGCGCGGCGCGCTGCTGACC 6546
Qy 826 oSerPro-----SerProSerProSerProSerAr 836
Db 6547 GGCACCGCGGAGCGGACAGTCAGGTGCGTATCGTGTGTCACGCGGCGACACCGCAACCG 6606
Qy 836 gSerProSerProSerAlaSerProSerProSerSer-----Se 849
Db 6607 ATCAGGTACCGTCCGCGGCGCGGCAACTTCAGCTTCGCGCGCGCGCGCTGATC 6666
Qy 849 rProSer-----ProSerSerProSerSerSe 859
Db 6667 ACCGCGCACTGATCGCGGCGGTTCCTGTCGACGCGCGCGCAACGTACGCGGCGCGCC 6726
Qy 859 rProSer-----ProThrPr 864
Db 6727 ACCATCAACGCGCGGACCTGGCGCGCGCGACCATCAGCGTGCAGGAGCGCGCATACC 6786
Qy 864 oSerSerProValSerGlyValLysValGlnTyrLysAsnAsnAspSerAlaPr 884
Db 6787 TGGATCAACGCGCGGAGATCGGGACGCGC-ATCCAGGTC----- 6825
Qy 884 oGlyAspAsnGlnIleLysProGlyLeuGlnValValAsnThrGlySerSerValAs 904
Db 6826 -----GATGTGACGCTCCGCGCATCCAGTCCAGTCCGCG- 6858
Qy 904 pLeuSerThrValThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuVa 924
Db 6859 -----CAGTGTGTCAAGTTC---GCCGCGCAGAACGCTTACGAGCGCGAGGTGAG 6911
Qy 924 lTyAsnCysAspTrpAlaAlaIleGlyCysGlyAsnIleArgAlaSerPhe----- 941
Db 6912 CCATACCTTCACCGCGCGGACATCGCGCGGCAACCTGACCTCCGCGCGCTCC 6971
Qy 942 -GlySerValAsnPro-----AlaThrProThrAlaAsp 952
Db 6972 CGCGGCGATGGCGCGTTCGCGGAGGCGCGCTCGACGCTCACCGCGCAC 7020

```

RESULT 11

US-09-410-551B-28
; Sequence 28, Application US/09410551B
; Patent No. 6503737
; GENERAL INFORMATION:
; APPLICANT: KOSAN BIOSCIENCES, Inc.
; APPLICANT: REEVES, CHRISTOPHER

```

; APPLICANT: CHU, DANIEL
; APPLICANT: KHOSLA, CHAITAN
; APPLICANT: SANTIL, DANIEL
; APPLICANT: WU, KAI
; TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
; FILE REFERENCE: 30062-20026.00
; CURRENT APPLICATION NUMBER: US/09/410,551B
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 60/139,650
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: US 60/123,810
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/102,748
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 4767
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
; OTHER INFORMATION: PKS synthase fragment
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(4766)
US-09-410-551B-28

Alignment Scores:
Pred. No.: 7,82e-06 Length: 4767
Score: 296.50 Matches: 233
Percent Similarity: 35.2% Conservative: 101
Best Local Similarity: 24.5% Mismatches: 397
Query Match: 5.8% Indels: 223
DB: 3 Gaps: 46

US-09-917-376-1 (1-957) x US-09-410-551B-28 (1-4767)
Qy 9 LeuThrMetArgSerArgArgLeuValSerLeuLeuAlaAlaThrAlaSerPheAlaVal 28
Db 1827 CTCAGCGAACACGAGACGCGCTGCGCGCTACCTACGCGCGCGCGCGGATATA 1886
Qy 29 AlaAlaAlaLeuGlyValLeuProIleAlaIleThrAlaSerProAlaHisAlaAlaThr 48
Db 1887 CGGGCTGTGGCATCCACGCTG-----GCGGTGACACGCTCGGT-GTTCGAGCACCGCGC 1939
Qy 49 ThrGlnProTyr----- 52
Db 1940 CGTACTCTCTGGAGATGACACCGTCACCGGACCGCGGTGACCGACCCAGGATCGGT 1999
Qy 53 -----ThrTrpSerAsnValAlaIleGlyGlyGly-----PheVal 65
Db 2000 TGTCTTTCGCGGAGGGGTGGCAGTGGCTGGGATGGGCGAGTGCACCTGCGCGATTCGTC 2059
Qy 66 AspGlyIleValPheAsnGly-----AlaProGlyIleLeuTyrValArgThrAsp 83
Db 2060 GGTGTGTTCGCGGAGCGATGCGCGGCGGTGTCGCGCGGTTCGCGAGTTCGT---GGA 2116
Qy 84 lIleGlyGlyMetTyrArgTrpAspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrp 103
Db 2117 CTGGGATCTGTTCAGGTTCGTGATGATCCGCGCGGTGG-TGGACCGGTGTATG---TGG 2172
Qy 104 ValGlyTrpAsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsn 123
Db 2173 TCCAGCCC-GCTTCTCTGGCGCATG-----ATGTTTCTCTGGCGCG----- 2213
Qy 124 ThrAsnLysValTrpAlaAlaValGlyMetTyrThrAsnSer---TrpAspProAsnAsp 142
Db 2214 -----GTGTGCGAGCGCGCGGTGTCGCGCGGATGCGGTGTATCGGCCATTCGCAG 2264
Qy 143 GlyAlaIleLeuArgSerSerAspGlnGlyAlaThrTrpGln----- 156

```

Db 2265 GGTGAGATCGCGCAGCTTGTTGTGGCGGTGCGGTGTCACTACGCGATCGCGCCGGATC 2324
Qy 157 IleThrProLeuProPheLeuLysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGluArg 176
Db 2325 GTGACCTTGGCGACGACGAGCGCTCGCGGGCGTGGCGGGCGCGGCGATGGCATCC 2384
Qy 177 LeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGly 196
Db 2385 GTGCGCCCTG---CCCGCGAGAGTGTGAGCTGTGTCAGCGGGCC----- 2426
Qy 197 LeuTyrArgSerThrAspSerGlyAlaThrTyrSerGlnMetThrAsnPheProAspVal 216
Db 2427 ---TGGATCGCGGCCCAACAGCGCCCGCTCCACCGTGTACGCGGCACCCCGGAAGCG 2483
Qy 217 GlyThrTyrIleAlaAsnProThrAspThrGly----- 228
Db 2484 GTGACCATGCTCTCACCGCTCATGAGCGCACAGGGGTGCGGTGCGCGGATCACCGTC 2543
Qy 229 ---TyrGlnSerAspIleGlnGlyValValTyrPheVal-----AlaPheAspLysSer 244
Db 2544 GACTATGCTTCGACACCCCGCAGCTGAGCTGATCGCGACGAACCTACTCGACATCACT 2603
Qy 245 SerSerSerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspProAsnAsn 264
Db 2604 AGCGACGACGCTCGACAGCCCGCTCGTGGCTGCTGCGACCGTGGACGGCACC--- 2660
Qy 265 ProValPheTyr-----SerArgAspGlyGlyAlaThrTyrGlnAlaValProGlyAla 282
Db 2661 -----TGGTTCGACAGCCCGCTGGACGGGAGTACTGGTACCAGGAACCTCGCGTAA 2711
Qy 283 ProThrGlyPheIleProHisLysGlyValPheAspProValAsnHisValLeuTyrIle 302
Db 2712 CCGGTGCGTTTCCACCCCGCGCTGAGCCAGTGTGAGCGCCGACACCGTGTTCGTC 2771
Qy 303 AlaThrSerAsnThrGlyGlyProTyrAspGlySerSerGlyAspValTyrLysPheSer 322
Db 2772 GAGTCAAGCCCGCCCGCGGTGTTGTCAGCGGATGAGCAGCATGTCGTCACGGTTGCC 2831
Qy 323 Val-----ThrSerGlyThrTyrThrArgIleSerProValProSerThrAspThr 339
Db 2832 ACGTGGTGTGTCAGCGGCGGCGCCACCGCATG-----CTCACCGCCCTG 2879
Qy 340 AlaAsnAspTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIle 359
Db 2880 GCACAGCGCTAT-----GTCCAGCGGTTCACCGTCGAC----- 2912
Qy 360 MetValAlaThrGlnIleSerTyrTyrProAspThrIleIlePheArgSerThrAspGly 379
Db 2913 -----TGCGCCGCCATCCTC----- 2927
Qy 380 GlyAlaThrTyrThrArgIleTyrAspThrTyrSerTyrProAsnArgSerLeuAtgTyr 399
Db 2928 GGCACACCAACACCGGGTACTGGACCTTCGACCTTCCACCTTCCACACCGCGGTAC 2987
Qy 400 ValLeuAspIleSerAlaGluProTyrLeuThrPheGlyValGlnProAsnProProVal 419
Db 2988 -----TGCTC-----GAGTGGCTCCCGCGCC 3011
Qy 420 ProSerProLysLeuGlyTyrMetAspGluAlaMetAlaIleAspProPheAsnSerAsp 439
Db 3012 ACGGCC-----GACTCGGC 3026
Qy 440 ArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTyrAsp 459
Db 3027 CACCCCGTCTCGGCACCGAGTGGCGTGGCGGG----- 3062
Qy 460 SerGlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluThrAla---Val 478
Db 3063 TCGCCGGCGGGTGTTCAGGGTCCCGTCCCGCGCGGTGCGGACCGCGGGTGTTCATC 3122
Qy 479 AsnAspLeuIleSerProProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGly 498
Db 3123 GCGGAACCTGGCGTTCGCGCGCGCGGACCC-----GACTGCGCC 3164

Qy 499 GlyPheThrHisAlaAspValThrAlaValProSerThrIlePheThrSerProValPhe 518
Db 3165 ACGGTGAAACAGCTCGAGCTCACCTCCGTCCGCGCGGATCCGCGCGGAGGCGC--- 3221
Qy 519 ThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGly 538
Db 3222 ACCGCGAGACTCGGTTCGAT----- 3242
Qy 539 SerPheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGly 558
Db 3243 -----GAAACCGCGCGCGCGCGCGCTTCACCGTCCACACCGCGCTCGCGAC 3296
Qy 559 LysAsnTyrPheGlnGlySerGluProGlyGlyValThrThrGlyGlyThrValAlaAla 578
Db 3297 GCCCGTGGAGCTGACGCGAG-----GGGGTTCCTCCGCGCGCGCGCTG---CC 3347
Qy 579 SerAlaAspGlySerArgPheValTyrAlaProGlyAspProGlyGlnProValTyr 598
Db 3348 CAGCCCCGAAGCGTGCACACCGCTGCGCCCG-----CCGGC-----CGGTGCGC 3395
Qy 599 AlaValGlyPheGlyAsnSerTyrAlaAlaSerGlnGlyValProAlaAsnAlaGlnIle 618
Db 3396 GCGACGGCTGCCCGGGCGTGGCGACGCGCGGACCGAGTCTTCGTGAACCCGAAGTC 3455
Qy 619 ArgSerAspArgValAsnProLysThrPheTyrAla-----LeuSerAsnGlyThr 635
Db 3456 GACAGC-----CCTGAGCGCTTCGTGGCAGACCCCGACCTGCTGCAGCGGTGTC 3503
Qy 636 PheTyrArgSerThrAspGlyGlyValThrPheGlnProValAlaAlaGlyLeuProSer 655
Db 3504 TTCTCCCGGTGCGCGACGGG-----AGCGCCAGCGCGACCGGATG---GCGCGACCTGCG 3556
Qy 656 SerGlyAlaValGlyValMetPheHisAlaValProGlyLysGluGlyAspLeuTyrLeu 675
Db 3557 GGTGACGCGTTCGACGCGCACCGTGTGCGCGCTGCTCACCCCGCGCAGTGGTGT 3616
Qy 676 AlaAlaSerSerGlyLeuTyrHisSerThrAsnGlyGlySerSerTyrSerAlaIleThr 695
Db 3617 CGTGGAGCTCGCGCCCTTCGAGCGGTGCGGATGCGGT-----GCTCACCGC 3664
Qy 696 GlyValSerSerAla-----ValAsnValGlyPheGlyLysSerAlaProGlySerSer 713
Db 3665 GGAGTGGTGAAGCTGCGCGAGGTGCGTTCGCGAGCGGATCCGACGAGTTCGACGGTCT 3724
Qy 714 TyrProAlaValPheValValGlyThrIleGlyGlyVal---ThrGlyAlaTyrArgSerAs 733
Db 3725 GCTTCGGCTTTCAGTGGTTCGCGGTGGCGGAGGCCCTACGACGCTGCC-----GA 3775
Qy 733 pAspCysGlyThrThrTyrValLeuIleAsn-----AspAspGlnHisGlnTyr 749
Db 3776 CGAGTCCCGAGGGGTACACCTTCATCCGCGCACACACCCCGACGACCCCGGACGACCC 3835
Qy 749 rGlyAsnTyrGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgVal----- 766
Db 3836 CACCAACCCCAACACACCCACGCGCACACACAAACACACACCGGTCTTCACCGC 3895
Qy 767 -----TyrIleGlyThrAsnGlyArgGlyIleValTyrGlyAspIle----- 780
Db 3896 CCTCAACACCCATCATCACCAACACACCTCATCTGCTCCACACACCCACCGACCC 3955
Qy 781 ---GlyValAlaProSerGly---SerProSerPro-----SerValSerProSerAlaAs 797
Db 3956 CCCAGGCGCGCGTTCACCGGCTTCACCGCACCGACCAACAAACGAAACCGCGCGCAT 4015
Qy 797 erProSerLeuSerPro-----SerProSerProSerSerSerProSer 811
Db 4016 CCACCTCATCGAAACCCACACACCCCGACACCGGCTCCCTCCCTACCCAACTCACCACT 4075
Qy 811 erProSerPro-----SerProSerSerSerProSerSerProSerProSerProSer 825
Db 4076 CCACCAACCCACCTACGCTTCACCAACACACCTTCACACACCCCGCCACCTCACCCCAT 4135

[illegible]

US-09-940-316B-28
; Sequence 28, Application US/09940316B
; Patent No. 6759536
; GENERAL INFORMATION:
; APPLICANT: KOSAN BIOSCIENCES, Inc.
; APPLICANT: REEVES, CHRISTOPHER
; APPLICANT: CHU, DANIEL
; APPLICANT: KOSLA, CHAITAN
; APPLICANT: SANTI, DANIEL
; APPLICANT: WU, KAI
; TITLE OF INVENTION: POLYPEPTIDES ENCODING THE FKBA GENE OF THE FK-520 POLYKETIDE SYNTHETASE
; TITLE OF INVENTION: GENE CLUSTER

```

Alignment Scores:
Pred. No.:      7-82e-06      Length:      4767
Score:          296.50         Matches:     233
Percent Similarity: 35.2%       Conservative: 101
Best Local Similarity: 24.5%    Mismatches:  397
Query Match:     5.8%          Indels:      223
DB:              3            Gaps:        46

US-09-917-376-1 (1-957) x US-09-940-316B-28 (1-4767)

Qy      9 LeuThrMetArgSerArgIguValSerLeuLeuAlaAlaThrAlaSerPheAlaVal 28
Db      1827 CTGACCGAACAAGAACGGCTGCCTACTTGGCGCGTCGCCCGGGCGGATATA 1886
Qy      29 AlaAlaAlaLeuClyValLeuProIIleThrAlaSerProAlaHisAlaAlaThr 48
Db      1887 CGGGCTGTGGCATCGACGCTG-----GCCGTGACACCGTTCGGT-GTTTCGAGCACCCGCGC 1939
Qy      49 ThrGlnProTyr----- 52
Db      1940 CGTACTCCTTGAGATGACACCGTCACCGGCACCGCGGTGACCGACCCCAGGATCGTGT 1999
Qy      53 -----ThrTrpSerAsnValAlaIIleGlyGlyGly-----PheVal 65

```


Db 2928 GGCACCAACACACCGGGTACTGGACCTTCGACCTACGCTTCCACACACGCGGTAC 2987
 QY 400 ValLeuAspIleSerAlaGluProTrpLeuThrPheGlyValGlnProAsnProProVal 419
 Db 2988 -----TGGCTC-----GAGTCGGCTCCCGGCC 3011
 QY 420 ProSerProLysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAsp 439
 Db 3012 ACGGCC -----GACTCGGC 3026
 QY 440 ArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAsp 459
 Db 3027 CACCGCTCTCGGACCGAGTCGCGTCCGCGG-----3062
 QY 460 SerGlyGlyGlnIleHisIleAlaProMetValLysLeuGluThrAla-----Val 478
 Db 3063 TCGCGGGCGGGTGTACGGGTCCCGTCCCGCGCGGTGCGGACCGCGGGTGTTCATC 3122
 QY 479 AsnAspLeuIleSerProProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGly 498
 Db 3123 GCGGAACCTGGCGTCCCGCGCGCGCGACCC-----GACTCGGC 3164
 QY 499 GlyPheThrHisAlaAspValThrAlaValProSerThrIlePheThrSerProValPhe 518
 Db 3165 ACGTTCGAACACTGACGTACCTCCGTCCCGCGGTTCGCGCGCGCGCGGCGC-----3221
 QY 519 ThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGly 538
 Db 3222 ACCGCGACACTGGGTGCAT-----3242
 QY 539 SerPheAspSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGly 558
 Db 3243 -----GAACCGCGCGCGCGCGCGCGCGCTTCACCGTCCACACCGCGTCCGCGC 3296
 QY 559 LysAsnTrpPheGlnGlySerGluProGlyGlyValThrThrGlyThrValAlaAla 578
 Db 3297 GCGCGTGCAGCTGACACCGCGAG-----GGGTTCGCGCGCGCGCGCGG-----CCC 3347
 QY 579 SerAlaAspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProValValTyr 598
 Db 3348 CAGCCGGAAGCGTGCACACCGCTGCGCCCG-----CGGCG-----GGGTGCC 3395
 QY 599 AlaValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIle 618
 Db 3396 GCGGACGGCTCCCGGGCGGTGGCGACGCGCGCGAGCTCTCTCGAAGCGGAGTC 3455
 QY 619 ArgSerAspArgValAsnProLysThrPheTyrAla-----LeuSerAsnGlyThr 635
 Db 3456 GACAGC-----CCTCAGCGCTTCGTGGCACACCGACCTGCTCGACGCGGTC 3503
 QY 636 PheTyrArgSerThrAspGlyGlyValThrPheGlnProValAlaAlaGlyLeuProSer 655
 Db 3504 TTCTCGCGGTGGGACCGG-----ACCGCGACCGACCGGATG-CGGGACCTCGC 3556
 QY 656 SerGlyAlaValGlyValMetPheHisAlaValProGlyLysGluGlyAspLeuTrpLeu 675
 Db 3557 GGTGCACGCGTGGGACCGCACCGTGTGCGCGCTCCCTACCGCGCGGACAGTGTGT 3616
 QY 676 AlaAlaSerSerGlyLeuTyrHisThrAsnGlyGlySerTrpSerAlaIleThr 695
 Db 3617 CGTGAGCTCGCGCTTCGAGGTGCGGATGCGGATGCGGT-----GCTACCGC 3664
 QY 696 GlyValSerSerAla-----ValAsnValGlyPheGlyLysSerAlaProGlySerSer 713
 Db 3665 GGAGTCGTGACGTGGGCGAGGTGCGTCCGCGGCGGCGGATCGGACGAGTCGCGGTCT 3724
 QY 714 TyrProAlaValPheValValGlyThrIleGlyGlyVal-ThrGlyValTyrArgSerAs 733
 Db 3725 GCTTCGGCTTGAGTGTTCGGTGGCGGAGGCCACTACGACGTGCC-----GA 3775
 QY 733 pAspCysGlyThrThrTrpValLeuIleAsn-----AspAspGlnHisGlnTyr 749
 Db 3776 CGAGTGGCGGAGGTACACCTCATCCCGGCCACACACCGCGGACCGCGACCGACCC 3835

QY 749 rGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgVal-----766
 Db 3836 CACCAACCCCAAC 3895
 QY 767 -----TyrIleGlyThrAsnGlyArgGlyIleValTyrGlyAspIle-----780
 Db 3896 CTTCAACACACACCTCATCCACCAACACACACACACACCTCATCGTCCACACACAC 3955
 QY 781 ----GlyGlyAlaProSerGly-SerProSerPro-----SerValSerProSerAlas 797
 Db 3956 CCAGGCGCGCGCTGACCGGCTTACCCGCGACCGACCAAAACGACACACCCCGCGCAT 4015
 QY 797 erProSerLeuSerPro-----SerProSerProSerSerSerProSerProSer 811
 Db 4016 CCACCTCATCGAAACCCACACCCACACCCACACCCACCTCCCTCCACCACTCACCCCT 4075
 QY 811 erProSerPro-----SerProSerProSerProSerProSerProSerProSer 825
 Db 4076 CCACCAACCCACCTACGCTTCACCAACACACCTCCACACCCCGCTCACCCCAT 4135
 QY 825 erProSerProSerProSerProSerProSerProSerProSerProSerProSer 845
 Db 4136 CACCAACCCACACACACACACACACACACACACCCCGACACCCCTCACCCCAACA 4195
 QY 845 erProSerSerSer-----ProSerProSerSerSerProSerProSerSerP 860
 Db 4196 CGCATCTCATCACCGCGGCTCCGCGACCTCCGCGGATCTCTCGCGGACCTCAA 4255
 QY 860 roSerProThrProSerSerProPro 868
 Db 4256 CCACCCCGACACCTACCTCTCTCC 4281
 RESULT 13
 US-09-410-551B-32
 ; Sequence 32, Application US/09410551B
 ; Patent No. 6503737
 ; GENERAL INFORMATION:
 ; APPLICANT: KOSAN BIOSCIENCES, Inc.
 ; APPLICANT: REEVES, CHRISTOPHER
 ; APPLICANT: CHU, DANIEL
 ; APPLICANT: KHOSLA, CHAITAN
 ; APPLICANT: SANTI, DANIEL
 ; APPLICANT: WU, KAI
 ; TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
 ; FILE REFERENCE: 30062-20026.00
 ; CURRENT APPLICATION NUMBER: US/09/410,551B
 ; PRIOR FILING DATE: 1999-10-01
 ; PRIOR APPLICATION NUMBER: US 60/139,650
 ; PRIOR FILING DATE: 1999-06-17
 ; PRIOR APPLICATION NUMBER: US 60/123,810
 ; PRIOR FILING DATE: 1999-03-11
 ; PRIOR APPLICATION NUMBER: US 60/102,748
 ; PRIOR FILING DATE: 1998-10-02
 ; NUMBER OF SEQ ID NOS: 72
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 32
 ; LENGTH: 4818
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
 ; OTHER INFORMATION: PKS synthase fragment
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (3)...(4817)
 US-09-410-551B-32
 Alignment Scores:
 Pred. No.: 1,18e-05 Length: 4818
 Score: 293.00 Matches: 235
 Percent Similarity: 33.1% Conservative: 97

Best Local Similarity:	23.4%	Mismatches:	379
Query Match:	5.7%	Indels:	295
DB:	3	Gaps:	48
US-09-917-376-1 (1-957) x US-09-410-551B-32 (1-4818)			
QY	1	MetAspArgSerGluAsnIleArgLeuThrMetArgSer	13
DB	1798	CTGGACCCGCTCCCGCGCGCGCGCTCAGCACCCGGCGAAGACCTTCGGCTGCTCGTGT	1857
QY	14	ArgArgLeuValSerLeuLeuAlaAlaThrAlaSerPheAlaValAlaAlaLeuGly	33
DB	1858	CGGCGGTTCCCGGAGCACTCGACGACAGATCGGCGCTCGCG	1905
QY	34	ValLeuProIleAlaIleThrAlaSerProAlaHisAlaAlaThrGlnProTyrThr	53
DB	1906	-----CCTATC-----TCGACACCGCGCGCGGCTCGACCGCGCGCGTGGCGC	1950
QY	54	-----TyrSerAsnValAlaIleGlyGlyGlyPheValAspGlyIleValPheAsn	71
DB	1951	AGACACTGG-----	1959
QY	72	GluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyr	88
DB	1960	-----CCGCGC-----GTACGCATTCACCCAGCGCGCGTACTGCTCGGGG	2001
QY	88	-----	88
DB	2002	ACACGTCATCGGCGCTCCCGCGCGGACAGCGCGCACTCGTCTTCGCTACTCCG	2061
QY	89	-----ArgTyrAspAla	92
DB	2062	GTACGGGCACCCAGCATCCCGCGATGGCGGAGCAGTAGCCGATTCGTGCGTGTGTCG	2121
QY	93	AlaAsnGly-----	95
DB	2122	CCGACCGGATGCGCGAGTGTGCGGCGGTTCGCGAGTTCGTGGACTCGTGTCA	2181
QY	96	-----ArgTyrIleProLeuLeuAspTyrValGlyTyrAsnAsnTyr	109
DB	2182	CGGTTCTGGATCATCCGCGGTGTGGACCGGTTGATGTGTCCAGCCC-GCTTCCCTGG	2240
QY	110	GlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTyrAla	129
DB	2241	GCATG-----ATGGTTTCTCTGCGCGC-----GTGTGGCAG	2273
QY	130	AlaValGlyMetTyrThrAsnSer-----TyrAspProAsnAspGlyAlaIleLeuArgSer	148
DB	2274	CGGCGCGGTGTGCGCGCGATGCGGTGATCGGCCATTCGCGAGGTGAGATCGCCGAGCT	2333
QY	149	SerAspGlnGlyAlaThrTyrGln-----IleThrProLeuProPhe	162
DB	2334	TGTGTGGCGGTGCGGTGTCTACTACGCGATGCGCGCGCGGATCGTGACCTTCGCGAGCCAG	2393
QY	163	LysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsn	182
DB	2394	CGCATCGCGCGCGCTGCGCGCGCGCGCGGATGCGATCGCTCGCCCTG-----CCGCGG	2450
QY	183	AsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGlyLeuTyrArgSerThrAsp	202
DB	2451	CAGGATGTCGAGCTGTGTCGCGCGGCGC-----TGATCGCGCGCCAC	2492
QY	203	SerGlyAlaThrTyrSerGlnMetThrAsnPheProAspValGlyTyrIleAlaAsn	222
DB	2493	AACGGCGCGCTCCACCGTATCGCGGCGCACCCCGGAAGCGGTGACCATGTCCTCACC	2552
QY	223	ProThrAspThrThrGly-----TyrGlnSerAspIle	233
DB	2553	GCTCATGAGGCACAGGGGTGCGGGTGCAGGATCACCGTGCAGTATGCCTCGCACACC	2612
QY	234	GlnGlyValValTyrVal-----AlaPheAspLysSerSerSerLeuGlyGln	250
DB	2613	CGGCACGTCGAGCTGATCCGCGGACGAACTACTCGACATCATCTAGCAGCAGCTCGCAG	2672

QY 605 SerTIPAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsn 624
Db 3465 GGTGGGACCGCGGACGAGTCTCGTCGAAGCCGAAGTCGACAGC----- 3512
QY 625 ProLysThrPheTyrAla-----LeuSerAsnGlyThrPheTyrArgSerThrAsp 641
Db 3513 CCTGACGGTTCGTGGACACACCCCGAGCTGCTCGACGGGTCTTCTCCGGGTCGGGAC 3572
QY 642 GlyGlyValThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyVal 661
Db 3573 GGG-----AGCGCCGACCGCGCGGATG-CGCGACCTCGCGGTGCACGGTCGGACGC 3625
QY 662 MetPheHisAlaValProGlyLysGlyGlyAspLeuThrPheLeuAlaAsnSerGlyLeu 681
Db 3626 CACGTGCTCGCGCTGCTCCACCGCGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3685
QY 682 TyrHisSerThrAsnGlyGlySerSerTyrPheAlaIleThrGlyValSerAla--- 700
Db 3686 CGACGGTCCGGATGCCGT-----GCTCACCGCGGAGTGGTGGCTGGG 3733
QY 701 ---ValAsnValGlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheVal 719
Db 3734 CGAGGTGCGTCCGCGAGCGATCCGACGAGTCGACGGTCTGCTTGGCTGTGAGTGT 3793
QY 720 ValGlyThrIleGlyGlyVal-ThrGlyAlaTyrArgSerAspCysGlyThrThr 739
Db 3794 GCCGGTGGCGGAGCCACTACGACGTGC-----GACGAGTCCCGGAGGCTA 3844
QY 739 pValLeuIleAsn-----AspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAl 755
Db 3845 CACCTCATCACCGCACACACCCCGGAGCCCGGAGCCACCCACCAACCCCAACAC 3904
QY 755 allerThrGlyAspHisAlaAsnLeuArgVal-----TyrIle 768
Db 3905 ACCCACGACCCACACACACACACCGCTCTCACCGCTCTCAACACCACTCAT 3964
QY 768 eGlyThrAsnGlyArgGlyIleValTyrGlyAspIle-----GlyGlyAlaProSe 785
Db 3965 CACCACCAACACACACCTCATCTGCTCCACCAACCAACCCCGCGCGCTCAC 4024
QY 785 rGly-SerProSerPro-----SerValSerProSerAlaSerProSerLeuSerPro- 802
Db 4025 CGSCCTCACCGCACCGCACAAACCAACACCCCGCGCGCTCATCCTCATCGAAACCCA 4084
QY 803 -----SerProSerProSerSerProSerProSerPro----- 814
Db 4085 CCACCCCCACACCCCACTCCCTCCACCACTCACCACTCACCACTCACCACTCAC 4144
QY 815 -----SerProSerSerSerSerSerSerSerProSerProSerProSerPro 831
Db 4145 CTTACCAACCAACACCTCCACACCCCGCGCTCACCCCTCACCACTCACCACTCAC 4204
QY 831 exProSerProSerArgSerProSerProSerProSerProSerProSerSerSer 849
Db 4205 CACCACCAACCCCAACACCCCGCTCAACCCCAACCCCGCTCACTCATCACCGG 4264
QY 850 -----ProSerProSerSerProSerProSerProSerProSerProSerPro 866
Db 4265 CGGCTCGGACCTCGCGGATCTCTGCGCGGACCTCAACCCCAACCCCACTTACCT 4324
QY 866 exSerPro 868
Db 4325 CCTCTCCC 4332

RESULT 14

US-09-940-316B-32
; Sequence 32, Application US/09940316B
; Patent No. 675936
; GENERAL INFORMATION:
; APPLICANT: KOSAN BIOSCIENCES, Inc.
; APPLICANT: REEVES, CHRISTOPHER
; APPLICANT: CHU, DANIEL
; APPLICANT: KHOSLA, CHAITAN

; APPLICANT: SANTI, DANIEL
; APPLICANT: WU, KAI
; TITLE OF INVENTION: POLYKETIDES ENCODING THE fkbA GENE OF THE FK-520 POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062-20026.11
; CURRENT APPLICATION NUMBER: US/09/940,316B
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 09/410,551
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 60/139,650
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: US 60/123,810
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/102,748
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 4818
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
; OTHER INFORMATION: PKS synthase fragment
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(4817)
US-09-940-316B-32

Alignment Scores:
Pred. No.: 1.18e-05 Length: 4818
Score: 293.00 Matches: 235
Percent Similarity: 33.1% Conservative: 97
Best Local Similarity: 23.4% Mismatches: 379
Query Match: 5.7% Indels: 295
DB: 3 Gaps: 48

US-09-917-376-1 (1-957) x US-09-940-316B-32 (1-4818)

QY 1 MetAspArgSerGluAsnIleArgLeuThrMetArgSer----- 13
Db 1798 CTGACCGCTCCCGCGCGCGCGCTCAGCAGCGGCGAAGACCTTCGCTGCTGCTGT 1857
QY 14 ArgArgLeuValSerLeuLeuAlaAlaThrAlaSerPheAlaValAlaAlaLeuGly 33
Db 1858 CGGCGGTTCCTCCGAGGACACTCGACGAGAGATCGGCGCTGCGCG----- 1905
QY 34 ValLeuProIleAlaIleThrAlaSerProAlaHisAlaAlaThrGlnProTyrThr 53
Db 1906 -----CCTATC-----TCGACACCGCGCGCGCGCTGACCGCGCGCGTGGCGC 1950
QY 54 -----TrpSerAsnValAlaIleGlyGlyGlyPheValAspGlyIleValPheAsn 71
Db 1951 AGACACTGG----- 1959
QY 72 GluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyr----- 88
Db 1960 -----CCCGC-----GTACGCACTTCAACCCCGCGCGCTGCTGCGGG 2001
QY 88 ----- 88
Db 2002 ACACCGTCATCGCGCTCCCCCGGAGCAGCGGAGCAGTCTGCTTCTACTCCG 2061
QY 89 -----ArgTrpAspAla 92
Db 2062 GTCAGGACACCCAGCATCCCGCATGGCGAGCAGTAGCCGATTCGTGCTGCTGCTG 2121
QY 93 AlaAsnGly----- 95
Db 2122 CCGAGCGGATGGCGAGTGTGCGCGCGGTTCGCGAGTTCGTGGACTGTTGTTCA 2181
QY 96 -----ArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrp 109
Db -----ArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrp 109

Db 2182 CGGTTCTGGATGATCCGGCGGTGGTACCGGGTTGATGTGGTCCAGCCC-GCTTCCTGG 2240
Qy 110 GlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAla 129
Db 2241 GCGATG-----ATGGTTTCTCCGCGCGG-----GTGTGCAG 2273
Qy 130 AlaValGlyMetTyrThrAsnSer---TrpAspProAsnAspGlyAlaIleLeuArgSer 148
Db 2274 GCGCGCGGTGTGGCGCGGATCGCGGTGATCGGCCATTGCGAGGTGAGATCGCGCAGCT 2333
Qy 149 SerAspGlnGlyAlaThrTrpGln-----IleThrProLeuProPhe 162
Db 2334 TGTGTGCGGGTGGGTGCTACTACGCGATGCCCGCGGATCGTGCCTTCCGCGAGCCAG 2393
Qy 163 LysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsn 182
Db 2394 GCGATCGCGCGGCGCTGGCGGGCGCGGCGCGATGCGATCGTCCGTCGCCCTG---CCGCGC 2450
Qy 183 AsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThrAsp 202
Db 2451 CAGGATGTCAGCTGGTTCGACGGGGCC-----TGATGCGCGCGCCAC 2492
Qy 203 SerGlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsn 222
Db 2493 AACGGCGCGCTCCACCGTATCGCGGGCACCCCGAAGCGGTGCGACCATGCTCTCACC 2552
Qy 223 ProThrAspThrThrGly-----TyrGlnSerAspIle 233
Db 2553 GCTCATGAGGCACAAAGGGTGGCGGTGGCGGATCACCGTGCATGCTTCGCACACC 2612
Qy 234 GlnGlyValValTrpVal-----AlaPheAspLysSerSerSerLeuGlyGln 250
Db 2613 CCGCACGTCGAGCTGATCCGCGCGAAGTACTCTGACATCACTAGCAGCAGCAGCTCGCAG 2672
Qy 251 AlaSerLysThrIlePheValGlyValAlaAspProAsnAsnProValPheTrp----- 268
Db 2673 ACCCGCTGTCGCGTGGCTGTCGACCGTGGACGGCACC-----TGGTGCAC 2720
Qy 269 SerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPheIlePro 288
Db 2721 AGCCCGCTGGAGGGAGTACTGTGTACGGAACTGCTGACCGTGCAGTGTTCACCCC 2780
Qy 289 HisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrAsnThrGly 308
Db 2781 GCCGTGACCATGTCAGGCCCGCGCGACACCGTGTTCGTCGAGTGCAGCGCGCGCCG 2840
Qy 309 GlyProTyrAspGlySerSerGlyAspValTrpLysPheSerVal-----ThrSer 325
Db 2841 GTGTTGTTGACGGCGATGGACGAGATGCTGTCACGGTTCGACGCTGCGTGCAGCAG 2900
Qy 326 GlyThrTrpThrArgIleSerProValProSerThrAspThrAlaAsnAspTyrPheGly 345
Db 2901 GCGACGCCACCCGGATG-----CTCACCGCTTCGGCAGGCGCTAT----- 2942
Qy 346 TyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIle 365
Db 2943 GTCCACGGCTCACCGTCGAC-----2963
Qy 366 SerTrpTrpProAspThrIleIlePheArgSerThrAspGlyAlaThrTrpThrArg 385
Db 2964 -----TGGCCCGCATCTC-----GGCACCAACCAACCCGG 2996
Qy 386 IleTrpAspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAla 405
Db 2997 GTACTGACCTTCGACCTACGCTTCCACACACCGCGGTAC-----3038
Qy 406 GluProTrpLeuThrPheGlyValGlnProAsnProProValProSerProLysLeuGly 425
Db 3039 -----TGGCTC-----GATCGGCTCCCCCGCCACCGCC 3068
Qy 426 TrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThr 445
Db 3069 -----GACTCGGGCCACCCCGTCTCTCGGCACC 3095

Qy 446 GlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlyGlnIleHis 465
Db 3096 GGAGTCGCGCTCGCGCGG-----TCGCGCGCGCGGTGTTTC 3131
Qy 466 IleAlaProMetValLysGlyLeuGluThrAla---ValAsnAspLeuIleSerPro 484
Db 3132 ACGGTCGCGTCCGCGCGCGGTGGACCGCGCGGTGTTTCATCGCGAACTGGCGTCCGCC 3191
Qy 485 ProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAsp 504
Db 3192 GCCCGCGAGCCACC-----GACTGCGCCACGCTCGAACAGCTCGAC 3233
Qy 505 ValThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSerVal 524
Db 3234 GTCACCTCGTGGCGCGCGGATCCCGCGCGCGAGGGCC---ACCGCGCAGACCTGGGTG 3290
Qy 525 AspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspProSerSer 544
Db 3291 GAT-----GAAACCGCGCC 3305
Qy 545 GlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGly 564
Db 3306 GACGGCGCGCGCTTCACCGTCCACACCGCGTGGCGAGCGCCCGTGGAGCTGCAC 3365
Qy 565 SerGluProGlyGlyValThrThrGlyGlyThrValAlaAlaSerAlaAspGlySerArg 584
Db 3366 GCCGAG-----GGGGTTCGCGCGCGCGCGTG---CCCCAGCCCGAGCGCTCGAC 3416
Qy 585 PheValTrpAlaProGlyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsn 604
Db 3417 ACCGCTGGCGCGC-----CCGGC-----GCGTGGCGCGAGCGGTGCGCGGG 3464
Qy 605 SerTrpAlaAlaSerGlnGlyValProAlaAlaGlnIleArgSerAspArgValAsn 624
Db 3465 GCGTGGCGAGCGCGGACCGAGTCTTCGTGAAGCCGAGTGCAGACG-----3512
Qy 625 ProLysThrPheTyrAla-----LeuSerAsnGlyThrPheTyrArgSerThrAsp 641
Db 3513 CCTGACGGTTCGTGGCACACCCCGACCTGTCGACGGGTCTTCTCCGCGGTGGCGAC 3572
Qy 642 GlyGlyValThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyVal 661
Db 3573 GGG-----AGCGCGCAGCGCGCGATG---GCGGACCTCGCGGTGCACGCTCGAGCGC 3625
Qy 662 MetPheHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeu 681
Db 3626 CACCGTGTGCGCGCTCCCTCACCCCGCGACAGTGTGTGTCGAGCTCGCGCCCTT 3685
Qy 682 TyrHisSerThrAsnGlyGlySerSerTrpSerAlaIleThrGlyValSerSerAla--- 700
Db 3686 CGAGGTGCGCGAATGCGGT-----GCTACCGCGAGTGGTGCAGCTGGG 3733
Qy 701 ---ValAsnValGlyPheGlySerAlaProGlySerSerTyrProAlaValPheVal 719
Db 3734 CGAGTTCGCGTTCGCGCGCGATCCGAGTCGAGTCGCTGCTTCGCTTGAAGTGGT 3793
Qy 720 ValGlyThrIleGlyGlyVal-ThrGlyAlaTyrArgSerAspCysGlyThrThrTr 739
Db 3794 GCGGTGCGCGAGGCGCCACTACGAGCTGCC-----GACGAGCTGCCCGAGGCGTA 3844
Qy 739 pValIleAsn-----AspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAl 755
Db 3845 CACCTTCATCACCGCCACACACCCGAGCGACCCCGACCCCAACCCCAACAC 3904
Qy 755 aileThrGlyAspHisAlaAsnLeuArgVal-----TyrIle 768
Db 3905 ACCCACGCGCACCCACACAAACACAGCGTCTCTACCGCCCTCCACACCACTCAT 3964
Qy 768 eGlyThrAsnGlyArgGlyIleValTyrGlyAspIle-----GlyGlyAlaProSe 785
Db 3965 CACCACCAACCAACCTCATCTGCTCCACACCAACCAACCCCGCGCGCGCTCAC 4024

Qy	785	gGly-SerProSerPro-----SerValSerProSerAlaSerProSerLeuSerPro-	802
		: : : :	
		: : : :	
Db	4025	CGGCGCTCACCGGCACGCAAAACAAACACCCCGCGCGCATCCACCTCATCGAAACCCCA	4084
		: : : :	
Qy	803	-----SerProSerProSerSerProSerProSerPro-----	814
		: : : :	
Db	4085	CCACCCOCCACACCCCACTCCCGCTCTCACCCCAACTCACCAACCGCTCCACCAACCCCACTTACG	4144
		: : : :	
Qy	815	-----SerProSerSerSerProSerSerProSerProSerProSerProS	831
		: : : :	
Db	4145	CCTCACCAACACACACCTCCACACCCCGCCACCTCACCCCATCACCAACCCACCAACACAC	4204
		: : : :	
Qy	831	erProSerProSerArgSerProSerProSerAlaSerProSerProSerSerSer-	849
		: : : :	
Db	4205	CACCAACACACCCCAACACCCCTCATACCCCAACCAACCGCATCTCATACCGG	4264
		: : : :	
Qy	850	-----ProSerProSerSerSerProSerSerProSerProSerProSerS	866
		: : : :	
Db	4265	CGGCTCGGCACCTCGCGCGGCATCTTCGCGCGCCACCTCAACCAACCCCAACCACTTACCT	4324
		: : : :	
Qy	866	erSerPro 868	
		: : : :	
Db	4325	CCTCTCCC 4332	
		: : : :	

RESULT 15
 US-09-252-991A-13773/c
 ; Sequence 13773, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 13773
 ; LENGTH: 4131
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-13773

[illegible]

			:::			:		:		: ::	
Db	3912	AACGTGTCCTTCACTCCAGGCATCCCGCTGCCGGATGGCAGCGTGCTCAACGTGGTGGCG	3853								
Qy	119	AlaaSpProIleAsnThrAsnLysValTrpAla-			:::					--AlaValGlyMet	133
Db	3852	CGCAGCCCAGAAGCATGTGACAGTCGGCCGGCGGTGATCACGTGTGATGCGTGGCCCCCG	3793								
Qy	134	TyrThrAsnSerTrpAspProAsnAspGlyAla--	IleLeuArgSerSerAspGlnGly	152							
Db	3792	GCGCGCGGTGATCGATCCGAGCAACGGCACGAGATAAGCGGTACC	CGGAGGCCGC	3733							
Qy	153	AlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArgGly	172								
Db	3732	GCGACGTGATCCTCACCGAT-	GCGCGCGCAACCCG	3697							
Qy	173	MecGlyCluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAlaPro	192								
Db	3696	ATTGGCGACGCCACCGCC-	3679								
Qy	193	SerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThrAsn	212								
Db	3678	GACGGCAGCGCAACTGGAGTTCAACCCG-	3649								
Qy	213	PheProaspValGlyThrTyrlleAlaAsnProThrAspThrThrGlyTyrGlnSerAsp	232								
Db	3648	-GGCACCCCGCTGGCCAACGGCACCGTGATCAAACGCC-	3613								
Qy	233	IleGlnGlyValValTrpValAlaPheApeLysSerSerSerSerLeuGlyGlnAlaSer	252								
Db	3612	-GTGGCCAGGACCGCGGCCGAATACACGAGCGTCCGCCAGC	3571								
Qy	253	LysThrIlePhe-	ValGlyValAlaAspProAsnAsnProValPhe	267							
Db	3570	GTACCGTGTGATGCCATCGCCCGCGCGCGGTGATCAATCCGAGCAACGAGTCTGC	3511								
Qy	268	TrpSer-	269								
Db	3510	ATCAGCGGTACGGCGAAAGCGGGCCACGGTGATCCTCAACGACGGCAACGCCAACCCG	3451								
Qy	270	-ArgAspGlyGlyAlaThrTrpGlnAlavalProGlyAlaPro-	283								
Db	3450	ATCGCCAGGTACCGCCGAGCGCACGGCAAGTGGGCTTTCACGCCGCCACCGCGTGTG	3391								
Qy	283	-	283								
Db	3390	GCCATGGCACGGTGATCAATGCGTGGCCAGGACGCCCGGCAACAACAGCATGCC	3331								
Qy	284	-ThrGlyPheIleProHisLysGlyValPheAspProValAsnHis	298								
Db	3330	ACCAGCGCCACCGTCGACTCGTGGCGCCAGCAGCCCGGTGATCGATCGAGCAACGGT	3271								
Qy	299	ValLeuTyrlleAlaThrSerAsnThrGlyGlyProTyr-	---AspGlySerSer	315							
Db	3270	AGCGTATCGCGGTACCGCGAGGTGTGGCCAGGTGATCCTACACCGCAACCGC	3211								
Qy	316	GlyAspValTrpLysPheSerValThr- SerGlyThrTrpThrArgileSerProval	334								
Db	3210	AAACCGATCGGCAGGTACCGCGCGATGGCAGCGCAACTGGAGTTCACGCCCGGCACG	3151								
Qy	335	Pro-	---SerThrAspThrAlaAsnAspTyrPhe	344							
Db	3150	CCGCTGTCCAATGGCACGGTGTCAATGCGGTGGCCAGGACGCTCGCGCAACACCGC	3091								
Qy	345	GlyTyr-SerGlyLeuThrIleAspArgGlnHisPro-	---	356							
Db	3090	GGCCCCGACGACACCGTGGATCGGTGGCGCGCGCGCGCGGTGATCGACCCGAGC	3031								
Qy	357	AsnThrIleMetValAlaThrGlnIleSerTrpTrpProAspThrIleIlePheArgSer	376								
Db	3030	AACGGCAGGTGATCCCGGTACCGCGGAAGCGGTGCGACCGTGTCTCTC-	2980								
Qy	377	ThrAspGlyGly-	AlaThrTrpThrArgileTrpAspThrPrthr	390							

```
Db 2979 ACCGATGGCGGCGCAACCCGATCGGCCAGCCAGCCGCGATGGCAGCGCAACTGGAGC 2920
Qy 391 SerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAla----- 405
Db 2919 TTCACCGCGGCGCAGCGCGTGGCCAAACCGCACCGGTGATCAATCGGTGGCCCGCAGGATCCG 2860
Qy 406 -----GluProTyrLeuThrPhe-----GlyValGlnProAsnProPro 418
Db 2859 GCCGGCAATACACGCGGCCAGCACCGACCGACCGGTGGACCGGTGGCCCGCCGCCACCCCG 2800
Qy 419 ValProSerProLysLeuGlyTyrMetAspGluAlaMetAlaIleAspProPheAsnSer 438
Db 2799 GTGGTCAACCGCAGCAACGCG----- 2779
Qy 439 AspArgMetLeuTyrGlyThr-----GlyAlaThrLeuTyrAlaThrAsnAspLeu 455
Db 2778 ---AGCGTGATCGCGGTACCGCGGAAGCGCGGCCAGCGGTGATCCTCACC----- 2731
Qy 456 ThrLysTrpAspSerGlyGlnIleHisIleAlaProMetValLysGlyLeuGlu 475
Db 2730 -----GACGGCGCGCGCAAC-----CCGATCGCGCAGGTCAACCGCCGAC 2692
Qy 476 ThrAlaValAsnAspLeuIleSerProProSerGlyAlaProLeu----- 490
Db 2691 GGCAGCGCACTGGAGCTTCACGCC-----GGCAGCGCGCTGGCCAAACGGCTCGGTG 2638
Qy 491 IleSerAlaLeuGly---AspLeuGlyGlyPheThrHisAlaAspVal---ThrAlaVal 508
Db 2637 ATCAATGGCTGGCCAGGAGCGCGCGGCAACACCGCGCGCCGCGCCAGCACCGGTG 2578
Qy 509 ProSerThrIlePheThrSerProValPhe-----ThrThrGlyThrSerValAsp--- 525
Db 2577 GACTCGGTAGTCCCGCGCACCCCGGTGCTCGATCCGAGCAACCGGTACGGTGATCAGCGGT 2518
Qy 526 TyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspProSerSerGln 545
Db 2517 ACCCGCAAGCGGGCGGCGCGGTGATCTCCACGACGCGCGCGCAACCCGATACGGCAG 2458
Qy 546 ProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTyrPheGlnGlySer 565
Db 2457 GCC-----ACCGCGGATGGCAGCGCAACTGG-----AGCTTC 2425
Qy 566 GluProGlyGlyValThrGlyGlyVal-----AlaAlaSerAlaAspGlySer 583
Db 2424 ACTCGGCGCACCGCTGACCAACCGCACCGGTGATCAATCGGTGGCCCGCAGCAGCGCC 2365
Qy 584 ArgPheValTyrAlaPro----- 589
Db 2364 GGCACACACGCGGTCCGGTCAGCACCAACAGTGGACGCGGTGGCCCGCCACCCCGGTG 2305
Qy 590 -----GlyAspProGlyGlnProValVal 597
Db 2304 ATCGACCGGAGCAATGTGTCAAACTCAGCGGCACCGCGCAACCCGCGTCCGGGTGATC 2245
Qy 598 TyrAlaValGlyPheGlyAsn-----SerTrp 606
Db 2244 CTCACCGATGGCAATGGCAACCCGATCGGCCAGACCTCGCCGCGAGTAGCGGCACTGG 2185
Qy 607 AlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLys 626
Db 2184 ACCTTCACACCGGGCACGCG----- 2164
Qy 627 ThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThrPhe 646
Db 2163 -----CTGGCCACACGGCAGCGGTGTCTCAACGCC-----GTGGCCCGAC 2128
Qy 647 GlnProValAla---AlaGlyLeuProSerSerGlyAlaValGlyValMetPheHisAla 665
Db 2127 GACCGCGCGGCAATACACGCGGTCCGGCCAGCACCGGTGGATACGCGTGGCGCGCGCC 2068
Qy 666 ValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThr 685
Db 2067 ACGCG-----GTGATCAATGCCAGC 2047
```

Search completed: March 2, 2006, 12:21:12
Job time : 512.215 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 2, 2006, 12:19:26 ; Search time 1796.03 Seconds
(without alignments)
4406.259 Million cell updates/sec

Title: US-09-917-376-1

Perfect score: 5134

Sequence: 1 MDRSENIRLTWRRRLVSL.....RASFGSVNPATPTADTVLQX 957

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlp
-Q-/abss/ABSSWEB.spool/US09917376/runat_02032006_091500_8223/app_query.fasta_1
-DB=Published Applications NA Main -QWTS=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs06p
-USER=US09917376 @CGN 1 1 3095 @runat_02032006_091500_8223 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5134	100.0	2869	3	US-09-917-376-2
2	5134	100.0	2869	5	US-10-155-400-2
3	2487	48.4	9025608	6	US-10-156-761-1
4	2478	48.3	2646	6	US-10-156-761-1845
5	1631.5	31.8	2517	5	US-10-026-994-4
6	1631.5	31.8	2710	5	US-10-026-994-1
7	1570	30.6	9025608	6	US-10-156-761-1
c					Sequence 1, Appli

Alignment Scores:
Pred. No.: 0
Score: 5134.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Length: 2869
Matches: 956
Conservative: 0
Mismatch: 0

ALIGNMENTS

RESULT 1

US-09-917-376-2

; Sequence 2, Application US/09917376

; Publication No. US20040038334A1

; GENERAL INFORMATION:

; APPLICANT: DING, SHI-YOU

; APPLICANT: ADNEY, WILLIAM S.

; APPLICANT: VINZANT, TODD B.

; APPLICANT: HIMMEL, MICHAEL E.

; TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS

; TITLE OF INVENTION: CELLULOLYTIC

; FILE REFERENCE: 40197.4US01

; CURRENT APPLICATION NUMBER: US/09/917,376

; CURRENT FILING DATE: 2001-07-28

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 2

; LENGTH: 2869

; TYPE: DNA

; ORGANISM: Acidothermus cellulolyticus

; FEATURE:

; NAME/KEY: modified_base

; LOCATION: (2869)

; OTHER INFORMATION: a, c, t, g, other or unknown

US-09-917-376-2

8	1442	28.1	2217	6	US-10-156-761-2561	Sequence 2561, Ap
9	1167.5	22.7	3668	3	US-09-927-827-21	Sequence 21, Appl
10	1125.5	21.9	2646	7	US-10-395-241-11	Sequence 11, Appl
11	1124	21.9	2481	7	US-10-395-241-17	Sequence 17, Appl
12	1117	21.8	2367	7	US-10-395-241-13	Sequence 13, Appl
13	940	18.3	5698	7	US-10-420-191-1	Sequence 1, Appli
14	823	16.0	3687	3	US-09-917-384-2	Sequence 2, Appli
15	823	16.0	3687	3	US-09-917-383-2	Sequence 2, Appli
16	702	13.7	2289	3	US-09-917-378-2	Sequence 2, Appli
17	392.5	7.6	1103	8	US-10-653-047-7511	Sequence 7511, Ap
18	365.5	7.1	651	9	US-10-450-763-20760	Sequence 20760, A
19	361	7.0	2223	6	US-10-156-761-550	Sequence 550, App
20	346.5	6.7	2706	9	US-10-939-262-11	Sequence 11, Appl
21	345	6.7	18596	3	US-09-880-107-1590	Sequence 1590, Ap
22	345	6.7	18596	3	US-09-967-768A-119	Sequence 119, App
23	345	6.7	18596	3	US-09-954-531-124	Sequence 124, App
24	345	6.7	18596	3	US-09-954-531-348	Sequence 348, App
25	345	6.7	18596	8	US-10-629-313-140	Sequence 140, App
26	345	6.7	18596	9	US-10-843-641A-1191	Sequence 1191, Ap
27	345	6.7	18596	9	US-10-843-641A-1415	Sequence 1415, Ap
28	345	6.7	18596	9	US-10-843-641A-6264	Sequence 6264, Ap
29	331	6.4	1887	9	US-10-939-262-13	Sequence 13, Appl
30	330.5	6.4	7407	6	US-10-246-330-3	Sequence 3, Appli
31	330.5	6.4	7407	7	US-10-282-122A-30151	Sequence 30151, A
32	319.5	6.2	9495	8	US-10-669-161-71	Sequence 71, Appl
33	299.5	5.8	210528	5	US-10-087-192-289	Sequence 289, App
34	297	5.8	32591	6	US-10-085-117-187	Sequence 187, App
35	296.5	5.8	4767	3	US-09-940-316B-28	Sequence 28, Appl
36	296	5.8	96602	6	US-10-085-117-61	Sequence 61, Appl
37	293	5.7	4818	3	US-09-940-316B-32	Sequence 32, Appl
38	292.5	5.7	185548	5	US-10-175-523-62	Sequence 62, Appl
39	292.5	5.7	185548	10	US-11-099-266-62	Sequence 62, Appl
40	290	5.6	1232	7	US-10-761-169-17	Sequence 17, Appl
41	289	5.6	324	7	US-10-282-122A-13296	Sequence 13296, A
42	288.5	5.6	4571	3	US-09-940-316B-18	Sequence 18, Appl
43	288	5.6	288	6	US-10-437-708-111	Sequence 111, App
44	288	5.6	288	6	US-10-395-402-111	Sequence 111, App
45	288	5.6	288	8	US-10-257-199-111	Sequence 111, App

Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0
US-09-917-376-1 (1-957) x US-09-917-376-2 (1-2869)			
QY	1	MetAspArgSerGluAsnIleArgLeuThrMetArgSerArgArgLeuValSerLeuLeu	20
DB	1	ATGGATCGTTCCGAGAACATCCGCTGACTATGAGATCACCAGCAGATTGGTATCATCTGCTC	60
QY	21	AlaAlaThrAlaSerPheAlaValAlaAlaLeuGlyValLeuProIleAlaIleThr	40
DB	61	GCAGCCACTGCTGCTTCCGCGTGGCCGCTCTGGGAGTTCTGCCCATCGCAGTAACG	120
QY	41	AlaSerProAlaHisAlaAlaThrThrGlnProTyrThrTrpSerAsnValAlaIleGly	60
DB	121	GCCTTCTCTCGCGCACCGCGGCACGACTCAGCGGTACACCTGGAGCAACGTCGCGATCGG	180
QY	61	GlyGlyGlyPheValAspGlyIleValPheAsnGluGlyAlaProGlyIleLeuTyrVal	80
DB	181	GCAGCGCGCTTTGTGACGCGGATCGTCTTCAATGAAGGTGCACCGGGAATTTCTGTACGTG	240
QY	81	ArgThrAspIleGlyGlyMetTyrArgTrpAspAlaAlaAsnGlyValTrpIleProLeu	100
DB	241	CGGACGGACATCGGGGATGTATCGATGGATGCGGCAACGGCGGTGGATCCCTCTT	300
QY	101	LeuAspTrpValGlyTrpAsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAsp	120
DB	301	CTGGATTGGGTGGATGMAACAAATGGGGGTACAAACGGCGCTCGTCAGCATTTGCGCGAC	360
QY	121	ProIleAsnThrAsnLysValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspPro	140
DB	361	CCGATCAATACTAAACAAGGTATGGCGCGCCGCTCGGAATGTACACCAACAGCTGGGACCA	420
QY	141	AsnAspGlyAlaIleLeuAspSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeu	160
DB	421	AACGACGGAGGATTTCCGCTCTGCTGTATCAGGGCGCAACGTGGCAATAACCGCCCTG	480
QY	161	ProPheLysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAsp	180
DB	481	CCGTTCAAGCTTTGGCGCAACATCCCGGGCGTGGATGGCGAGCGCTTGGCGTGGAT	540
QY	181	ProAsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSer	200
DB	541	CCAAACAATGACAAACATTTCTGATTTCCGCGCCCGAGCGGCAAAAGGCTCTCGAGAAGC	600
QY	201	ThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyrIle	220
DB	601	ACGATTTCCGCGCGACCTGCTCCAGATGACGAACCTTTCCGAGACGTAGGCACGATCAT	660
QY	221	AlaAsnProThrAspThrThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAla	240
DB	661	GCAATCCCACTGACACACCGGTATCAGAGCGATATTAAGGGCTGCTCTGGGTGCT	720
QY	241	PheAspLysSerSerSerLeuGlyGlnAlaSerLysThrIlePheValGlyValAla	260
DB	721	TTCCACAAGTCTTCGTCATCGCTCGGCAACGAGTAAGACCAATTTTGTGGCGGTGGCG	780
QY	261	AspProAsnAsnProValPheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValPro	280
DB	781	GATCCCAATAATCCGGTCTTCTGGAGCAGACAGCGCGCGAGCTTGGCAGCGGTGCGG	840
QY	281	GlyAlaProThrGlyPheIleProHisLysGlyValPheAspProValAsnHisValLeu	300
DB	841	GGTGGCGGACCGGCTTCATCCCGCAAGGGCGTCTTTGACCGGTTCAACACGCTGCTC	900
QY	301	TyrIleAlaThrSerAsnThrGlyGlyProTyrAspGlySerSerGlyAspValTrpLys	320
DB	901	TATATTGCCACCAATACGGGTGGTCCGTATCAGCGGAGCTCGGCGAGCTCTGGAA	960
QY	321	PheSerValThrSerGlyThrTrpThrArgIleSerProValProSerThrAspThrAla	340
DB	961	TTCTCGGTGATCTCCGGGACATGACCGCAATCAGCCCGGTACCTTCGACGACACGCGCC	1020

QY	341	AsnAspTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMet	360
DB	1021	AACGACTACTTTTGGTTACAGCGGCTCACTATCGACCGCGACACCCGCAACAGATAATG	1080
QY	361	ValAlaThrGlnIleSerTrpTrpProAspThrIleIlePheArgSerThrAspGlyGly	380
DB	1081	GTGGCAACCCAGATATCGTGGTGGCGGACCAATAATTTTCGGAGCACCGACGGCGGT	1140
QY	381	AlaThrTrpThrArgIleTrpAspTrpThrSerTyrProAsnArgSerLeuArgTyrVal	400
DB	1141	CGACGCTGGACCGGATCTGGGATTTGGAGGTATCCCAATCGAAGCTTCGATATGTG	1200
QY	401	LeuAspIleSerAlaGluProTrpLeuThrPheGlyValGlnProAsnProValPro	420
DB	1201	CTTGACATTTCCGGGAGCTTGGCTGACCTTCGGGTACAGCGGAATCTCTCCGTACCC	1260
QY	421	SerProLysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArg	440
DB	1261	AGTCCGAAGCTCGGCTGGATGGATGGAATGGAATCGATCCGTTCAACTCTGTATCGG	1320
QY	441	MetLeuTyrGlyThrGlyValaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSer	460
DB	1321	ATGCTCTACGGAACAGGCGGACGTTGTACGCAACAATGATCTCACGAAGTGGGACTCC	1380
QY	461	GlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluThrAlaValAsnAsp	480
DB	1381	GGCGCCAGATTCATATCGCGCGATGTCAAAGGATTTGGAGAGACGCGGTAAACGAT	1440
QY	481	LeuIleSerProProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyPhe	500
DB	1441	CTCATCAGCCCGCGCTCTGGCGCCGCTCATCAGCGCTCTCGAGACCTCGCGCGCTTC	1500
QY	501	ThrHisAlaAspValThrAlaValProSerThrIlePheThrSerProValPheThrThr	520
DB	1501	ACCACGCGAGCTTACTGCGGTGCGCATCGAGATCTTCACGTACCGGTGTTACGACC	1560
QY	521	GlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPhe	540
DB	1561	GGCACCGGTGCGACTATGCGGAATTTGAATCCGTGATCATCTTCGCGCTGGAAGTTTC	1620
QY	541	AspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsn	560
DB	1621	GATCCATCGAGCAACCCGAAACGACAGGCACTGCGGTCTTCGACAGACGCGCGCAAGAAC	1680
QY	561	TrpPheGlnGlySerGluProGlyValThrThrGlyGlyThrValAlaAlaSerAla	580
DB	1681	TGTTTCCAAAGCAGCAACCTGGCGGGGTGACGAGCGGCGGCGGCTCGCCGATCGGCC	1740
QY	581	AspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProValValTrpAlaVal	600
DB	1741	GACGGCTCTCGTTTCTGCTGCGGTCCCGCGATCCCGGTACGCTGTGTGTGTACGCGATC	1800
QY	601	GlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSer	620
DB	1801	GGATTTGGCAACTCTCGGTGCTTTCGCAAGGTGTTCCGCCCAATGCCAGATCCGCTCA	1860
QY	621	AspArgValAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThr	640
DB	1861	GACCGGTGAATCCAAAGACTTTCTATGCCCTATCCAATGGAACCTTCTATCGAAGCAG	1920
QY	641	AspGlyGlyValThrPheGlnProValAlaAlaGlyLeuProSerSerGlyValAlaValGly	660
DB	1921	GACGCGCGGTGACATTCGAACCGGTGCGGCGGTCTTCGAGAGCGCGGTGCTCGGTGCT	1980
QY	661	ValMetPheHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGly	680
DB	1981	GTGATGTTCCACGCGGTGCTTGGAAAAGAGCGGATCTGTGGCTGCTGCTGATCGAGCGG	2040
QY	681	LeuTyrHisSerThrAsnGlyGlySerSerTrpSerAlaIleThrGlyValSerAla	700
DB	2041	CTTTACCACTCAACCAATGGCGGAGAGTTGGTCTGCAATCACCAGGCTATCTCCCGG	2100
QY	701	ValAsnValGlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheValVal	720


```
Db 2101 GTGAACGTGGGATTGGTAACTCGCGCCGGTCTCATACCGACCGCTTTGTCTCTC 2160
Qy GlyThrIleGlyGlyValThrGlyAlaTyrArgSerAspCysGlyThrTrpVal 740
Db 2161 GGCACGATCGGAGCGTTACGGGGCGGTACCGTCCGACGACTGTGGACGACCTGGGTA 2220
Qy LeuIleAsnAspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHis 760
Db 2221 CTGATCAATGATGACCAACCAATACGGAATTTGGGACAAAGCAATCACCGGTGACCAC 2280
Qy AlaAsnLeuArgArgValTyrIleGlyThrAsnGlyArgGlyIleValTyrGlyAspIle 780
Db 2281 GCGAATTTACCGCGGGTGTACATAGCAGCAAGCGCGTGGAAATGTATACGGGACAAAT 2340
Qy GlyGlyAlaProSerGlySerProSerProSerValSerProSerAlaSerProSerLeu 800
Db 2341 GGTGTGCGCGCTCGGATCGCGTCTCGTCCGTGAGTCCGTCCGCTTCGCCGAGCCGTG 2400
Qy SerProSerProSerProSerProSerProSerProSerProSerProSerProSerPro 820
Db 2401 AGCCGAGCCGAGCCGAGCAGCTCGGCATCGCGTCCGCGTCCGCGAGCTCGAGTCCA 2460
Qy SerSerSerProSerProSerProSerProSerProSerProSerProSerProSerPro 840
Db 2461 TCCTCGTCCGCGCTTCCTCGTCCGCGTCACCATCGCGAGTCCGCTCTCCGTCCACCA 2520
Qy SerAlaSerProSerProSerProSerProSerProSerProSerProSerProSerPro 860
Db 2521 TCGGGCTCGCGAGCCCGTCTTCGTACCGAGCCCGTCTTCGTACCCGCTCTTCGTCCG 2580
Qy SerProThrProSerSerSerProValSerGlyGlyValIleValIleValIleValIle 880
Db 2581 AGCCCAACGCGCTCGTCTCGCGGTGTCGGGTGGGTGAGGTGACGATATAAGAAATAT 2640
Qy AspSerAlaProGlyAspAsnGlnIleLysProGlyLeuGlnValValAsnThrGlySer 900
Db 2641 GATTCCGCGCGCGGTGATAATCAGATCAAGCGGGTTTGCAGGTGGTGAATACCGGGTGC 2700
Qy SerSerValAspLeuSerThrValThrValArgTyrTrpPheThrArgAspGlyGlySer 920
Db 2701 TCGTCCGTTGGATTGTCCAGCGGTGACGGTGGGTACTCGTTTACCCCGGATGGTGGCTCG 2760
Qy SerThrLeuValTyrAsnCysAspTrpAlaIleGlyCysGlyAsnIleArgAlaSer 940
Db 2761 TCGACACTGGGTGACAACTGTGACTGGGGCGGATCGGGTGTGGGAATATCCGGCCCTCG 2820
Qy PheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 956
Db 2821 TTCGGCTCGGTGAACCCGCGAGCCGCGAGCGGACACCTTACTTCGAG 2868
```

RESULT 2

```
US-10-155-400-2
; Sequence 2, Application US/10155400
; Publication No. US20030108988A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU.
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOYTICUS
; FILE REFERENCE: NREL 01-36A
; CURRENT APPLICATION NUMBER: US/10/155,400
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2869
; TYPE: DNA
; ORGANISM: Acidothermus cellulolyticus
; FEATURE:
; NAME/KEY: modified_base
```

```
; LOCATION: (2869)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-10-155-400-2
```

Alignment Scores:

```
Pred. No.: 0 Length: 2869
Score: 5134.00 Matches: 956
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0
```

US-09-917-376-1 (1-957) x US-10-155-400-2 (1-2869)

```
Qy 1 MetAspArgSerGluAsnIleArgLeuThrMetArgSerArgArgLeuValSerLeuLeu 20
Db 1 ATGGATCGTTCGGAGAACATCGCTCTGACTATGAGATCAGCAGATTGGTATCACTGCTC 60
Qy 21 AlaAlaThrAlaSerPheAlaValAlaAlaLeuGlyValLeuProIleAlaIleThr 40
Db 61 GCGCGCACTGCGTCTGTCGCCGTGGCGCGCTCTGGAGATTCTGCCCATCGCGATAACG 120
Qy 41 AlaSerProAlaHisAlaAlaThrThrGlnProTyrThrTrpSerAsnValAlaIleGly 60
Db 121 GCTTCTCTCGGCACGCGCGCAGACTCAGCCGTACACCTGGAGCAACGTGGCGATCGGG 180
Qy 61 GlyGlyGlyPheValAspGlyIleValPheAsnGluGlyAlaProGlyIleLeuTyrVal 80
Db 181 GCGCGCGCTTTGTTCGACGGGATCGTCTTCATGAGGTGCACCGGGAATCTGTACGTG 240
Qy 81 ArgThrAspIleGlyGlyMetTyrArgTrpAspAlaAlaAsnGlyArgTrpIleProLeu 100
Db 241 CGGACGCGACATCGGGGGATGTATCGATGGGATCGCCCAACGCGCGGTGGATCCCTCTT 300
Qy 101 LeuAspTrpValGlyTrpAsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAsp 120
Db 301 CTGGATTGGGTGGATGGAACAAATTTGGGGGTACAACGCGCTCTGTACAGCATTTGGCGG 360
Qy 121 ProfileAsnThrAsnLysValTrpAlaValGlyMetTyrThrAsnSerTrpAspPro 140
Db 361 CGGATCAATACTAAACAGGTATGGCGCGCTCGGAATGTACACCAACAGCTGGGACCCA 420
Qy 141 AsnAspGlyAlaIleLeuArgSerSerAspGlnGlyValAlaThrTrpGlnIleThrProLeu 160
Db 421 AACGACGCGAGCGATTCTCGCTCTGTCTGATCAGGGCGCAACGTGGCAATAAACGCCCTG 480
Qy 161 ProphelLysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAsp 180
Db 481 CCGTTCAAGCTTGGCGGCAACATGCGCGCGCTGGAAATGGCGGAGCGGCTTGGCGTGGAT 540
Qy 181 ProAsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyGlyGlyLeuTrpArgSer 200
Db 541 CCAAAACAATGACAACATTCTGTATTTCGGCGCGCCCGAGCGGCAAGGGCTCTCGGAGAGC 600
Qy 201 ThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyrIle 220
Db 601 ACAGATTCCGGCGCGACCTCGTCCAGATCAGCAACTTTCCGGACGTAGCAGCATATT 660
Qy 221 AlaAsnProThrAspThrThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAla 240
Db 661 GCAATCCCACTGACACGACCGGCTATCAGAGCATATTCAAGGCGTCTCTGGTGGCT 720
Qy 241 PheAspLysSerSerSerSerLeuGlyGlnAlaSerLysThrIlePheValGlyValAla 260
Db 721 TTCGACAAAGTCTTCGTCTCGTCCGGCAAGGAGTAAGACCAATTTTGTGGCGGTGGCG 780
Qy 261 AspProAsnAsnProValPheTrpSerArgAspGlyGlyValAlaThrTrpGlnAlaValPro 280
Db 781 GATCCCAATATCCCGTCTTCTGGAGCAGAGACGCGCGCGGCGGCGGCGGCGGCGGCGG 840
Qy 281 GlyAlaProThrGlyPheIleProHisLysGlyValPheAspProValAsnHisValLeu 300
Db 841 GGTGGCGCGACCGGCTTTCATCCCGCAAGGGCGTCTTTGACCCCGGTCAACACGCTGCTC 900
```

Qy 301 TyrileAlaThrSerAsnThrGlyGlyProTyrAspGlySerSerGlyAspValTrpLys 320
Db 901 TATATTGCCACACAGCAATACGGGTGGTCCGTATACGGGAGCTCCGGCAGCTCGGAAA 960
Qy 321 PheSerValThrSerGlyThrTrpThrArgIleSerProValProSerThrAspThrAla 340
Db 961 TTCCTGGTGACTCCGGGACATGACCGGAATACGCCGGTACCTTCGACGGACACGGCC 1020
Qy 341 AsnAspTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMet 360
Db 1021 AACGACTACTTTGGTTACAGCGGCTCCTATACGCCCGCAGCACCGACACCATATG 1080
Qy 361 ValAlaThrGlnIleSerTrpTrpProAspThrIleIlePheArgSerThrAspGlyGly 380
Db 1081 GTGGCAACCCAGATATCGTGTGGCCGACACCAATAATCTTCGGAGCACCGACGGCGT 1140
Qy 381 AlaThrTrpThrArgIleTyrAspThrTrpSerTyrProAsnArgSerLeuArgTyrVal 400
Db 1141 GCGACGTGGACCGGATCTGGGATGGACGAGTTATCCCAATCGAAGCTTCGCGATATGTG 1200
Qy 401 LeuAspIleSerAlaGluProTrpLeuThrPheGlyValGlnProAsnProProValPro 420
Db 1201 CTTGACATTTCCGGGAGCCTTGGCTGACCTTCGGCGTACAGCGAATCTCCCGTACCC 1260
Qy 421 SerProLysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArg 440
Db 1261 AGTCCGAAGCTCGGCTGGATGAAGCGATGGAATCGATCGCTTCAACTCTGATCGG 1320
Qy 441 MetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSer 460
Db 1321 ATGCTCTACCGGAACAGCGCGGACGTTGTACGCAACAAATATCTCACGAAGTGGGACTCC 1380
Qy 461 GlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluThrAlaValAsnAsp 480
Db 1381 GCGCGGCAGATTCTATATCGCGCCGATGGTCAAGGATTGGAGGAGCGCGGTAAACGAT 1440
Qy 481 LeuIleSerProProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPhe 500
Db 1441 CTCATCAGCCCGCGCTTGGCGCCCGCTCATCAGCGCTCTCGGAGACCTCCGGCGGCTC 1500
Qy 501 ThrHisAlaSerValThrAlaValProSerThrIlePheThrSerProValPheThrThr 520
Db 1501 ACCCACGCCAGCGTTACTGCGGTGCCATCAGACGATCTTCACGTCACCGGTGTTCCAGC 1560
Qy 521 GlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPhe 540
Db 1561 GGCACCGCGTGGACTATGCGGAATTGAATCCGTCGATCATCGTTCGCGTGGGAAGTTTC 1620
Qy 541 AspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsn 560
Db 1621 GATCCATCGAGCCNACCGAACGACGACGCGTCCGCTTCGACAGACGGCGGCGCAAGAC 1680
Qy 561 TrpPheGlnGlySerGluProGlyGlyValThrThrGlyThrValAlaAlaSerAla 580
Db 1681 TGGTTCCAAAGCAGCAACCTGGCGGGGTGACAGCGGGCGCACCGTCGCGCATCGGCC 1740
Qy 581 AspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProValValTyrAlaVal 600
Db 1741 GACGGCTCTCGTTTCGTGGGTCCCGGCGATCCCGGTGAGCTGTCGCTGTGAGCGATC 1800
Qy 601 GlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSer 620
Db 1801 GGATTTGGCAACTCTCGGTGCTTCCAAAGGTGTTCCGCCAATGCCAGATCCCGTCA 1860
Qy 621 AspArgValAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThr 640
Db 1861 GACCGGTGAATCCAAAGACTTCTATGCGCTATCCATGCAATGGAACCTTCTATCGAAGACG 1920
Qy 641 AspGlyGlyValThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGly 660
Db 1921 GACGGCGCGTGCATTTCCAAACCGGTTCGCGCGGTCTTCGAGCAGCGGTGCGCTCGGT 1980

Qy 661 ValMetPheHisAlaValProGlyLysGlyAspLeuTrpLeuAlaAlaSerSerGly 680
Db 1981 GTCATTTGCCACCGGGTGCCTGGAAGAGAGCGCATCTGTGGCTCGTCATCGAGCGGG 2040
Qy 681 LeuTyrHisSerThrAsnGlyGlySerSerTrpSerAlaIleThrGlyValSerAla 700
Db 2041 CTTTACCACACTCAACATATGGCGGACAGTTGGTCTGCAATCACCGGCTATCTCCCGG 2100
Qy 701 ValAsnValGlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheValVal 720
Db 2101 GTGAACGTGGGATTTGGTAAGTCTCGGCCCGGGTCTCATACCCAGCGCTCTTGTCTGTC 2160
Qy 721 GlyThrIleGlyGlyValThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpVal 740
Db 2161 GGCACGATCGGAGCGTTACGGGGCGGTACCGCTCCGACGACTGTGGGACGACCTGGGTA 2220
Qy 741 LeuIleAsnAspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHis 760
Db 2221 CTGATCAATGATGACCAATACGCAATATGGGACCAAGCAATCACCGGTACCCAC 2280
Qy 761 AlaAsnLeuArgArgValTyrIleGlyThrAsnGlyArgGlyIleValTyrGlyAspIle 780
Db 2281 GCGAATTTACGGCGGGTGTATACAGGACAGAACCGCGTGGAAATGTATATACGGGACATT 2340
Qy 781 GlyGlyAlaProSerGlySerProSerProSerValSerProSerAlaSerProSerLeu 800
Db 2341 GGTGGTGGCGCTCCGGATCCGCTCCGTCGCTGAGTCCGCTTCGCGCTTCGCGAGCCTG 2400
Qy 801 SerProSerProSerProSerProSerProSerProSerProSerProSerProSerPro 820
Db 2401 AGCCGAGCGCGACCGCGAGCAGCTCCGCTCCGCTGAGTCCGCTTCGCGAGCTCGA 2460
Qy 821 SerSerSerProSerProSerProSerProSerProSerProSerProSerProSerPro 840
Db 2461 TCCTGTCGCGCTCTCCGTCGCCGTACCATCGCGAGTCCGCTCTCGGTCTCCGTCACCA 2520
Qy 841 SerAlaSerProSerProSerProSerProSerProSerProSerProSerProSerPro 860
Db 2521 TCGGCGTCCGCGAGCGCTCTTCGTCACCGAGCGCTCTTCGTCACCGCTCTTCGTCGCG 2580
Qy 861 SerProThrProSerSerSerProValSerGlyGlyValLysValGlnTyrLysAsnAsn 880
Db 2581 AGCCCAACGCGCTCGTCCGCTCGCGGTGGGTGGAAGGTGACAGTATAGAAATAAT 2640
Qy 881 AspSerAlaProGlyAsnAsnGlnIleIleProGlyLeuGlnValValAsnThrGlySer 900
Db 2641 GATTCCGGCGCGGTGATAATCAGATCAAGCGGGTTTCAGAGGTGGTGAATACCGGGTGC 2700
Qy 901 SerSerValAspLeuSerThrValThrValArgTyrTrpPheThrArgAspGlyGlySer 920
Db 2701 TCGTCCGTGGATTTGTGACCGGTGACGGTGGGTACTGGTTCAACCGGGATGGTGGCTCG 2760
Qy 921 SerThrLeuValTyrAsnCysAspTrpAlaAlaIleGlyCysGlyAsnIleArgAlaSer 940
Db 2761 TCGACACTGGTGTAACAATGTGACTGGCGCGCATCGGTGTGGGAATATCCGCGCTCG 2820
Qy 941 PheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 956
Db 2821 TTCGGCTCGTGAAACCGCGGACCGCCGCGGACACCTACCTGCAG 2868

RESULT 3

US-10-156-761-1
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262
 CURRENT APPLICATION NUMBER: US/10/156,761
 CURRENT FILING DATE: 2002-05-29
 PRIOR APPLICATION NUMBER: JP 2001-204089
 PRIOR FILING DATE: 2001-05-30
 PRIOR APPLICATION NUMBER: JP 2001-272697
 PRIOR FILING DATE: 2001-08-02
 NUMBER OF SEQ ID NOS: 15109
 SEQ ID NO 1
 LENGTH: 9025608
 TYPE: DNA
 ORGANISM: Streptomyces avermitilis
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (4187715)
 OTHER INFORMATION: a, t, c, g, other or unknown
 US-10-156-761-1

Alignment Scores:
 Pred. No.: 3.4e-161 Length: 9025608
 Score: 2487.00 Matches: 490
 Percent Similarity: 64.8% Conservative: 135
 Best Local Similarity: 50.8% Mismatches: 248
 Query Match: 48.4% Indels: 93
 Ds: 6 Gaps: 18

US-09-917-376-1 (1-957) x US-10-156-761-1 (1-9025608)

QY 9 LeuThrMetArgSerArgLeuValSerLeuLeuAlaAlaThrAlaSerPheAlaVal 28
 DB 2277388 CTCACGGGCGAAGAACCCGATCTTC-----ACGGCGGTCTCGCGCTG 2277432
 QY 29 AlaAlaAlaLeu-----GlyValLeuProIleAlaIleThrAlaSerProAlaHisAla 46
 DB 2277433 GCGGCGGTCTGCCGCGCGGACCCCGCGACCTGGCGCGGAGCGCCCGACGCGGACG 2277492
 QY 47 AlaThrThrGlnProTyrThrTyrSerAsnValAlaIleGlyGlyGlyPheValAsp 66
 DB 2277493 ATCGCGCGGACAGTACGCTGGAAGAACCGCGCGGTGCGGCGGCTTCTGCTCCC 2277552
 QY 67 GlyIleValPheAsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGly 86
 DB 2277553 GGCATCGCTTCAACCGCTCCGGAAGAACCTCGCTACGCGCGGACCGACATCGCGGCG 2277612
 QY 87 MetTyrArgTyrAspAlaAlaAsnGlyArgTyrIleProLeuLeuAspTyrValGlyTyr 106
 DB 2277613 GCCTACCGCTGGCGGAGTCTCTCGAAGACCTCGGACCGCTGCTCGACTCGGCTGCGTGG 2277672
 QY 107 AsnAsnTyrGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLys 126
 DB 2277673 AGCGACTGGGGGCACACGGGTGTCTGAGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 2277732
 QY 127 ValTyrAlaAlaValGlyMetTyrThrAsnSerTyrAspProAsnAspGlyAlaIleLeu 146
 DB 2277733 GTGTACGCGGCGGTGCGGACGTACACCAAGTGGGACCGCGGCAACGGTGGCGGTGCTC 2277792
 QY 147 ArgSerSerAspGlnGlyAlaThrTyrGlnIleThrProLeuProPheLeuGlyGly 166
 DB 2277793 AGTCCGCGGACCGGGGCGGAGTGGCAGAAAGACCGGACCTGCGCTTCAAGCTCGGCGGG 2277852
 QY 167 AsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIle 186
 DB 2277853 AACATGCGGCGCGGGGATGGGCGAGCGGTCTCGGCTCGACCCGACGAGAACAGCGGTG 2277912
 QY 187 LeuTyrPheGlyAlaProSerGlyGlyLeuTyrArgSerThrAspSerGlyAlaThr 206
 DB 2277913 CTGTATCTCGGCGGCGGACGCGGCAAGGGGTGTGGCGGTGCGGACTCGGCGGCGCTCC 2277972
 QY 207 TrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThr 226
 DB 2277973 TGGTCGAGGTACCGACTTCCCGAACGCTCGGCACCTACGTACGAGGACCGGACGACG 2278032
 QY 227 ThrGlyTyrGlnSerAspIleGlnGlyValIleThrValAlaPheAspLysSerSer 246

DB 2278033 AGCGGTACGGTCCGACCAACAGGCGCATCTGTGGTCACTTCACGAGTCCAGCGGG 2278092
 QY 247 SerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspProAsnAspVal 266
 DB 2278093 TCGCGGGGAGCTCCACGCGGAGCGGTGTACGTGGGTGCGGAGCAAGCAACTCCGTC 2278152
 QY 267 PheTyrSerArgAspGlyGlyAlaThrTyrGlnAlaValProGlyAlaProThrGlyPhe 286
 DB 2278153 TATCGCTCCAGGACGCGGCGCGACCTGTGCTCCGCGCGGCGGACCGCCAT 2278212
 QY 287 IleProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsn 306
 DB 2278213 CTCGCCCAAGAGCGGTCTGTGACGCGGGAACGGCTGTGTGTACCTCGGTACGCGAC 2278272
 QY 307 ThrGlyGlyProTyrAspGlySerSerGlyAspValTyrLysPheSerValThrSerGly 326
 DB 2278273 AAGGCGGACCGTACGACGCGGCAAGGACAGCTGTGGCGGTACACGACGAGACCGGG 2278332
 QY 327 ThrTyrThrArgIleSerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyr 346
 DB 2278333 ACCTGGACGAACATCAGCCCGGTGCGGAGGCGGACACAC-----TACTACGGCTTC 2278383
 QY 347 SerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSer 366
 DB 2278384 AGCGGCTGACGTCGTGACCGCGGACGATCGGCGGACGTCGTGCGTACGACTCC 2278443
 QY 367 TrpTyrProAspThrIleIlePheArgSerThrAspGlyGlyAlaThrTyrPheArgIle 386
 DB 2278444 TGGTGGCGGACACGACGCTCTTCGCTCCACGACGCGGCGGACCTGGACGAGGCC 2278503
 QY 387 TrpAspTyrThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGlu 406
 DB 2278504 TGGGACTACACTGTATCCGAGCGCTCGAACCGCTTCCACCTGATGTCTGCTCTCG 2278563
 QY 407 ProTyrLeuThrPheGlyValGlnProAsnProProValProSerProLysLeuGlyTyr 426
 DB 2278564 CCCTGGCTCACCTGGGAGCGAACCAGCGCGGACGACAGCCCGGAACTCGGCTGG 2278623
 QY 427 MetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThrGly 446
 DB 2278624 ATGACCGAGTCCCTGGAGATCGACCGCTTCGACCGCGGATGATGACGGAACCGGG 2278683
 QY 447 AlaThrLeuTyrAlaThrAsnAspLeuThrLysTyrAspSerGlyGlyGlnIleHisIle 466
 DB 2278684 GCGACGCTACGCGACCGACCACTGACGAAGTGGGACGCGGAGACCGACGTCACATC 2278743
 QY 467 AlaProMetValLysGlyLeuGluThrAlaValAlaAsnAspLeuIleSerProProSer 486
 DB 2278744 AAGCGATGGCGGGGCGCTGGAGGAGACGCGCTCAACGCTCGCTCGCTCCCTCC 2278803
 QY 487 ---GlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspVal 505
 DB 2278804 GCGGCGGCGGAGCTGTTCAGCGCGCTCGGTGACATCGGGCGGTTCGCGGACACGACCT 2278863
 QY 506 ThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSerValAsp 525
 DB 2278864 ACCGCGTGGCTGCTGATGTACAGCTCGCGGAACTTCCACGACGACGACCTCGAC 2278923
 QY 526 TyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspProSerSerGln 545
 DB 2278924 TACGCGGAGACCGACCGCGGCGGTCGTGGGTGCGGCTCGGCAATCTCGAC-----TCGGGT 2278977
 QY 546 ProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnThrPheGlnGlySer 565
 DB 2278978 CCG-----CATGTGGCTTCTCGACGGAACACGCGGCGCAACTGTGTCGCGGGGCG 2279028
 QY 566 GluProGlyGlyValThrGlyThrValAlaAlaSerAlaAspGlySerArgPhe 585
 DB 2279029 GACCCCTCGGGGTGCGGCGGTCGCGGCGGTCGCGGCGGTCGACGCGGCTCGCTTC 2279088
 QY 586 ValTyrAlaProGlyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSer 605

```
Db 2279089 GTGTGGAGCCCGGGCCACCGGG-----GTGCAGTACACGACCGGGTTCGGACCTCG 2279142
Qy 606 TrpAlaalaSerGlnGlyValProAlaAlaAlaGlnIleArgSerAspArgValAsnPro 625
Db 2279143 TGGTGGCGGTCCGGGGCGCTCCGGCGGGGGGATCTCGAGTCCGACCGGGTCGACCG 2279202
Qy 626 LysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyValThr 645
Db 2279203 AAGACCTTCACGGCTTCAAGTCCGGACGGTTCACGTCTGAGTTCGGACGGGGCGGAC 2279262
Qy 646 PheGlnProValAlaAla---GlyLeuProSerSerGlyAlaValGlyValMetPheHis 664
Db 2279263 TTCACGGCTCCGGGGCCACGGGCTCCGGAGCGCGCACAGC-----GTGCGTTCAG 2279316
Qy 665 AlaValProGlyLysGluGlyAspLeuTrpLeuAlaalaSerSer----- 679
Db 2279317 GCGCTGCGCGCACGAGGGCGACATCTGGCTGGCGGGCGGCGGAGCGGCGGTAC 2279376
Qy 680 GlyLeuTyrHisSerThrAsnGlyLysSerSerTrpSerAlaIleThrGlyValSerSer 699
Db 2279377 GGGCTGTGGCACTCGACGAGCGGGCGGCGCTTCACCAAGCTCGCCACCGTCGAC 2279436
Qy 700 AlaValAsnValGlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheVal 719
Db 2279437 GCGACACCATCGCTTCGGCAAGCGGCGGCGCTTCACCAAGCTCGCCACCGTCGAC 2279496
Qy 720 ValGlyThrIleGlyValThrGlyValTyrArgSerAspCysGlyThrThrTrp 739
Db 2279497 AGCGGAAGATCGCGGTGTGCGCGCATCTTCGGTCGACCGACAGGGCGGAGCTGG 2279556
Qy 740 ValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAsp 759
Db 2279557 ACCCGGCTCAACGACGATGCGCCACGACGAGTGGGCTGGAGCGGCGGCGCATCAG 2279616
Qy 760 HisAlaAsnLeuArgArgValTyrIleGlyThrAsnGlyArgGlyIleValTyrGlyAsp 779
Db 2279617 CCACGGGTACGGGCGCGGTATGTGTGACGAGACGGGCGGGATGCTACGGCGAC 2279676
Qy 780 IleGlyGlyAlaProSerGly-----SerProSerProSerValSerProSerAla 796
Db 2279677 ACCCGGGCTTCGAGCGGGCGGCTACGGAGCGGCGCACCGAGCGGCGCTGCACGGTG 2279736
Qy 797 SerProSerLeuSer-ProSerProSerProSerProSerProSerProSerProSe 815
Db 2279737 ACATACAGGATCAGAACACGAGTGTGCGGCGGCTTCAGGCGCATGATGATGATG 2279796
Qy 815 rProSerSerSerProSerSerSer-----Pr 824
Db 2279797 ACCGGGTGACCGCTGGGACGGCTGCTGCTGGGCTGCTGCTGCTGCTGCTGCTGCT 2279856
Qy 824 oSerProSer----- 827
Db 2279857 GTACCCAGCTGTGGAACGGCTGTACGCGAGCGCGGTTGCGGGGTGACCGAGCGAAC 2279916
Qy 828 -----ProSerPro----- 830
Db 2279917 CTGGCTTGAACCGGAGGGTGGCGGCGGCTCTCGGTGAGCTTCGGGTTCACGGGAAGC 2279976
Qy 831 -----SerProSerProSerProSerProSerProSerProSerProSerProSe 843
Db 2279977 TGGTCGGGATCCAAATGGCAGACCGACCGCTTCAAACCTCGCGGCTCAGAACTGCA 2280036
Qy 843 rPro-----SerProSerProSerProSerProSerProSerProSerProSe 861
Db 2280037 GCCTGATGCAAGCCGCA-CGCTCGCTGATTCGAGGCGGCGGCGGCGGCGGCGGCGGCG 2280095
Qy 861 rProThrProSerSerSerProValSerGlyValLysValGlnTyrLysAsnAsnAs 881
Db 2280096 CCCTCGGCTCTCTCCACGGGTGTAGACCGTGGCTG-----GCAA 2280137
Qy 881 pSer-AlaProGlyAspAsnGlnIleLysProGlyLeuGlnValValAsnThrGlySer 901
Db 2280138 CTCCCGTGTGCTGAGGCTTCGCGCAGTTCGCGGCTCCGCGTGGGTGCGGCGGCGAGTA 2280197
```

```
Qy 901 erSerValAsp 904
Db 2280198 CCCGGCACGAC 2280208
```

RESULT 4

```
US-10-156-761-1845
; Sequence 1845, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1845
; LENGTH: 2646
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2646)
US-10-156-761-1845
```

```
Alignment Scores:
Pred. No.: 2,97e-164 Length: 2646
Score: 2478.00 Matches: 473
Percent Similarity: 64.1% Conservative: 128
Best Local Similarity: 50.4% Mismatches: 227
Query Match: 48.3% Indels: 110
DB: 6 Gaps: 14
```

US-09-917-376-1 (1-957) x US-10-156-761-1845 (1-2646)

```
Qy 23 ThrAlaSerPheAlaValAlaAlaAlaLeu-----GlyValLeuProIleAlaIleThr 40
Db 22 ACGCCCGTGTCTGCGCTGGCGCGGCTCTGCCCGCGCACCCCGCGGCACTGGCCGCG 81
Qy 41 AlaSerProAlaHisAlaAlaThrThrGlnProTyrThrTrpSerAsnValAlaIleGly 60
Db 82 AGCGCCCGCCGCGGACGATCGCGCGGACAGTACAGTACAGTGGAGAACGCCCGGCTCGAC 141
Qy 61 GlyGlyGlyPheValAspGlyIleValPheAsnGluGlyAlaProGlyIleLeuTyrVal 80
Db 142 GCGCGCGCTTCGTCGCCGCGCATCGTCTTCAACCGCTCCGAGAACACCTCGCCTACGCC 201
Qy 81 ArgThrAspIleGlyGlyMetTyrArgTrpAspAlaAlaAsnGlyArgTrpIleProLeu 100
Db 202 CGCACCGACATCGCGCGGCGCTACCGCTGGCGCGGAGTCTCGAAGACCTGGACGCGCTG 261
Qy 101 LeuAspTrpValGlyTrpAsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAsp 120
Db 262 CTCAGTCGGTGGCTGAGGACTGGGGGACACGCGGTGCTGAGGCTTCGCCCTCCGAC 321
Qy 121 ProIleAsnThrAsnLysValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspPro 140
Db 322 TCCGTCCAGCCGACAAAGGTGACGCGCGCTCGGCAGCTACACGAAACAGCTGGGACCGC 381
Qy 141 AsnAspGlyAlaIleLeuArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeu 160
Db 382 GGCAACGGTGGCGTCTCAGGTCCGCGGACCGCGGCGGAGTGGCAGAGACCGACCTG 441
```

QY 161 ProPheLeuGlyGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAsp 180
DB 442 CCTTTCAAGCTGGCGGGAACAATGCCGGCCCGGGCATGGCGGAGCGGCTCGCGTGCAC 501
QY 181 ProAsnAsnAsnAlaLeuTyrPheGlyAlaProSerGlyGlyLeuTyrArgSer 200
DB 502 CCGAACAGACAGCGGTCTGTATCTCGCGGCGCCAGCGCAAGGGCTGTGGCGGTGC 561
QY 201 ThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyrIle 220
DB 562 ACGGACTCGGGGCTCTCTGTGTCGAGGTCAACGACTTCCCGAACGTCGGCACTACGTG 621
QY 221 AlaAsnProThrAspThrThrGlyTyrGlnSerAspIleGlnGlyValValTyrValala 240
DB 622 CAGGACGCGACCGACACAGCGGGTACGCTCCGACCAACAGGGCATCTGTGTGGGTCA 681
QY 241 PheAspLysSerSerSerLeuGlyGlnAlaSerLysThrIlePheValGlyValala 260
DB 682 TTCGACGAGTTCACCGGGTCCCGGGAGCTCCACGCGGACGGTGTACGTCTGGGTTCGC 741
QY 261 AspProAsnAsnProValPheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValPro 280
DB 742 GACAGGACAACTCCGTATCGCTCCACGAGCGCGGCGGACCTGGTCCCGGTGGCC 801
QY 281 GlyAlaProThrGlyPheIleProHisLysGlyValPheAspProValAsnHisValLeu 300
DB 802 GCGCAGCCACCGGCATCTCGCCCAACAGGCGGTCTGGACCGCGCAACGGCTGTCTG 861
QY 301 TyrIleAlaThrSerAsnThrGlyGlyProTyrAspGlySerSerGlyAspValTrpLys 320
DB 862 TACCTCGGTACAGCAAGGGCGGACCGTACGACCGCGGACGAGCAGCTGTGGCGG 921
QY 321 PheSerValThrSerGlyThrTrpThrArgIleSerProValProSerThrAspThrAla 340
DB 922 TACACAGACACCGGACCTGGAGCAATACACCCGCTCCGAGGCGCGACAC--- 978
QY 341 AsnAspTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMet 360
DB 979 -----TACTACGGCTTCAGCGGGCTGACCGTGGACCGGACGATCCGGGACGGTGATG 1032
QY 361 ValAlaThrGlnIleSerTrpProAspThrIleIlePheArgSerThrAspGlyGly 380
DB 1033 GCGACTCGGTACAGCTCTCTGTGGCGGACACGACGCTCTTCGCTCCACGACGACGCGC 1092
QY 381 AlaThrTrpThrArgIleThrAspTrpThrSerTyrProAsnArgSerLeuArgTyrVal 400
DB 1093 GGCACCTGGACGAAGCGCTGGGACTACACCTCGTATCCGAGCGCTCCGAACCGCTTCAC 1152
QY 401 LeuAspIleSerAlaGluProTrpLeuThrPheGlyValGlnProAsnProValPro 420
DB 1153 ATGGATGTCTGTCTCGCCCTGGCTACCTGGGAGCGGAACCCGCGACCGCCGAGCAG 1212
QY 421 SerProLysLeuGlyTyrTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArg 440
DB 1213 ACCCGMAACTCGGTGTGATGATGACCGAGTCCCTGGAGATCGACCCGTTCGACTCCGCGCG 1272
QY 441 MetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSer 460
DB 1273 ATGATGTACGGAACGGCGCGGACGCTACGCGACGCGACACCTGACGAACCTGGGACACG 1332
QY 461 GlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluThrAlaValAsnAsp 480
DB 1333 GGAAGCCAGTTCCATCAAGCCGATGGCGGGCGCTGGAGGAGCGCGCTCAACGAC 1392
QY 481 LeuIleSerProProSer---GlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGly 499
DB 1393 CTCGCTCGCTCCCTCCGCGGCGCCCGAGCTGTTCAGGCGCTCGGTGACATCGCGCGC 1452
QY 500 PheThrHisAlaAspValThrAlaValProSerThrIlePheThrSerProValPheThr 519
DB 1453 TTCGCGCACACGACCTCACACGCTGGCGTGTGATGATACACGTCGCGCAACTTCACC 1512
QY 520 ThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySer 539

DB 1513 ACAGACACAGGCTCGACTACCGCGGAGCCGCCGGGACCGTGTGTGGGTTCGGCAAT 1572
QY 540 PheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLys 559
DB 1573 CTCGAC-----TCGGGTCCG-----CATGTGGGTTCGACGAGCAACGCGGCC 1617
QY 560 AsnTrpPheGlnGlySerGluProGlyGlyValThrThrGlyGlyThrValAlaAlaSer 579
DB 1618 AACTGTGTTCCGGGGCGGACCTTCGCGGGTCAGCGGGGTGGACGCTCGCGGCGCGC 1677
QY 580 AlaAspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProValValTyrAla 599
DB 1678 TCCGACGCGAGTCGCTTCGTGTGGAGCCCGCGGCGCACCGGG-----GTGCAGTACACG 1731
QY 600 ValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArg 619
DB 1732 ACCGGTTTCGGACACTTCGTGTGGGTCCGGGGCTCCCGCGCGGGGCGATCGTCGAG 1791
QY 620 SerAspArgValAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSer 639
DB 1792 TCCGACCGGGTCGACCCGGAAGACCTTCTACGGCTTCAAGTCCGCGAGGTTCTACGTCA 1851
QY 640 ThrAspGlyGlyValThrPheGlnProValAlaAla---GlyLeuProSerSerGlyAla 658
DB 1852 TCGGACGCGGGCGGACCTTCACGGCTCCGGCGCCACCGGCGCTCGGACGCGCGACG 1911
QY 659 ValGlyValMetPheHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSer 678
DB 1912 -----GTCCGCTTCAAGCGCTGCCCGGACGAGAGGGCGACATCTGGCTGGCGGCGCGC 1965
QY 679 Ser-----GlyLeuTyrHisSerThrAsnGlyGlySerSerTrpSerAla 693
DB 1966 GCGAGCGGCGCGCTACCGGCTGTGGCACTCGACGACGCGCGCGCGGCTTCCACCAAG 2025
QY 694 IleThrGlyValSerSerAlaValAsnValGlyPheGlyLysSerAlaProGlySerSer 713
DB 2026 CTCGCGCACCGTCGACCGAGCCGACCATCGGCTTCGGCAAGCGCGACCGCGCTCG 2085
QY 714 TyrProAlaValPheValValGlyThrIleGlyGlyValThrGlyAlaTyrArgSerAsp 733
DB 2086 TACCAAGACCTCTACACCGCGGAGATCGCGGCTGTGGCGGCACTTCCGGTCCACC 2145
QY 734 AspCysGlyThrThrTrpValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnTrpGly 753
DB 2146 GACAGGGCGCGAGCTGGACCGGCTCAACGACGATGCCACCGCTGGGGTTGACGCGGC 2205
QY 754 GlnAlaIleThrGlyAspHisAlaAsnLeuArgArgValTyrIleGlyThrAsnGlyArg 773
DB 2206 GCGGCGATCACCGGTGACCCCGGCTACGGGCGCGTGTATGTGTGCGACGAGCGGCGC 2265
QY 774 GlyIleValTyrGlyAspIleGlyAlaProSerGlySerProSerProSerValSer 793
DB 2266 GGGATCGTCTACGGCGACACCGCGGCTCTCGGACGCGCGC----- 2307
QY 794 ProSerAlaSerProSerLeuSerProSerProSerProSerProSerProSerProSer 813
DB 2307 ----- 2307
QY 814 ProSerProSerSerProSerProSerProSerProSerProSerProSerProSer 833
DB 2307 ----- 2307
QY 834 ProSerArgSerProSerProSerAlaSerProSerProSerProSerProSerProSer 853
DB 2307 ----- 2307
QY 854 SerSerProSerSerSerProSerProSerProSerProSerProSerGlyVal 873
DB 2308 -----GGTACGAGCGCGCACCGACG-----GGCGCTGC 2337
QY 874 LysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGlyLeu 893
DB 2337 ----- 2337

```

Db 2338 ACGGTGACATACAGGATCAGAACACAGTGGTGGCGGCTTCAG-----GCCGATGTA 2391
Qy 894 GlnValValAlaThrGlySerSerValAspLeuSerThrValThrValArgTyrTrp 913
Db 2392 CAGCTCGCCCAACACCGGCTGACCGCTGGGACGGCTGGTGGTGGG-----TGG 2442
Qy 914 PheThrArgAspGlySerSerThrLeuValTyrAsnCysAspTrpAlaAlaIleGly 933
Db 2443 TCGTTCGGTGACGGGACAGGAGGTACCCAGCTGTGGAACGCTCGTACGGCAGCGCGT 2502
Qy 934 CysGly-----AsnIleArgAlaSerPheGlySer 943
Db 2503 TCGGGGTGACCGCAGCGAACCTGGCTGGACCGGAGGTGGCGCGGCTCC 2556

RESULT 5
US-10-026-994-4
; Sequence 4, Application US/10026994
; Publication No. US20030113732A1
; GENERAL INFORMATION:
; APPLICANT: Dunn-Coleman, Nigel
; APPLICANT: Goedegebuur, Frits
; APPLICANT: Ward, Michael
; APPLICANT: Yao, Jian
; TITLE OF INVENTION: EGVI Endoglucanase and Nucleic Acids
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: GC698
; CURRENT APPLICATION NUMBER: US/10/026,994
; CURRENT FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 2517
; TYPE: DNA
; ORGANISM: Trichoderma reesei
US-10-026-994-4

Alignment Scores:
Pred. No.: 7,53e-105 Length: 2517
Score: 1631.50 Matches: 353
Percent Similarity: 57.2% Conservative: 130
Best Local Similarity: 41.8% Mismatches: 297
Query Match: 31.8% Indels: 64
DB: 5 Gaps: 23

US-09-917-376-1 (1-957) x US-10-026-994-4 (1-2517)

Qy 25 SerPheAlaValAlaAlaLeuGlyValLeuProIleAlaIleThrAlaSerProAla 44
Db 10 TCTCGAGTCTTGGCTTGGCTTGGCGGCGGTC-----ATCCCTGCC 51
Qy 45 HisAlaAlaThrThrGlnProTyrThrTrpSerAsnValAlaIle---GlyGlyGlyGly 63
Db 52 CATGCTGCC-----TTTTCATGAAGACGCTCAAGCTCGCGCGCGCGCGG 99
Qy 64 PheValAspGlyIleValPheAsnGluGlyAlaProGlyIleLeuTyrValArgThrAsp 83
Db 100 TTGCTCCCGCATCATCTTCATCCACAGCAAAAGCGGTAGCATATGACGACAGAT 159
Qy 84 IleGlyMetTyrArgTrpAspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrp 103
Db 160 ATTGGGGGCTGTACCGCTCAAC---GCCGACGACTCATGGACCGCGCTCACGGATGGG 216
Qy 104 Val-----GlyTyrAsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAla 119
Db 217 ATTGCTGATTAATGCCGCTCGCAACTGG-----GGCATCGACGCTGTGGCGCTT 267
Qy 120 AspProIleAsnThrAsnLysValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAsp 139
Db 268 GATCCGACGACCATCAAAAGGTATGGCCAGTCCGATGATATAGAAACAGCTGGGAT 327
Qy 140 ProAsnAspGlyAlaIleLeuArgSerAspGlnGlyAlaThrTrpGlnIleThrPro 159
Db 328 CCGAGTAATGGAGCCATCATCTCGTCTGTCAGACCGCGGCGCAACGTTCTTCACCAAC 387

```

```

Qy 160 LeuProPheLysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaVal 179
Db 388 TTGGCTTCAAAGTCGGGGTAAATCATGCAGACGCGAGCGGAGCGTCTGGCTGTC 447
Qy 180 AspProAsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArg 199
Db 448 GATCGGCACTCAACATCATCTACTTTGGTCTCGCTCAGGAAACGGCTCTGGAAG 507
Qy 200 SerThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyr 219
Db 508 TCTACGAGCGCGGCTGACCTTTTCAGGTCTGCTGTTTCCAGCAACTCGGCACTGGAGTAC 567
Qy 220 IleAlaAsnProThrAspThrThrGlyTyrGlnSerAspIleGlnGlyValValTrpVal 239
Db 568 ATCCAGACCGGAGTGATTCACACGGCTACAACAGCGACAGCAAGAGGACTCATGTGGT 627
Qy 240 AlaPheAspLysSerSerSerLeuGlyGlnAlaSerLysThrIlePheValGlyVal 259
Db 628 ACGTTCGACTCAACAGCAGCAGCAGCGGCGGAGCCAGTCTCGTATCTTTTGGCAGC 687
Qy 260 AlaAspPro---AsnAsnProValPheTrpSerArgAspGlyGlyAlaThrTrpGlnAla 278
Db 688 GCTGATAACATCACTGCTTCACTCTATGTGACAGCAATGCCGCTCCACGTGGAGTGT 747
Qy 279 ValProGlyAlaProThrGlyPheIleProHisLysGlyValPheAspProValAsnHis 298
Db 748 GTACCGGGCAGCAGGGAATACTTCTCACAAGGCGCAACTCAGCCAGCAGAGAG 807
Qy 299 ValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyrAspGlySerGlyAspVal 318
Db 808 GCCTTGATCTGACCTATTCCGATGGCAGCGGCGGTATGATGGCACACTTGGCTCAGTG 867
Qy 319 TrpLysPheSerValThrSerGlyThrTrpArgIleSerProValProSerThrAsp 338
Db 868 TGGAGGTACGACATTGCGGGGAACTTGGAAAGACATCACCCCTGTCTCTGGATCAGAT 927
Qy 339 ThrAlaAsnAspTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThr 358
Db 928 CTA-----TACTTTGGCTTGGCGGCTTGGCTCGATTTGCAAAAGCAGCAAC 978
Qy 359 IleMetValAlaThrGlnIleSerTrpTrpProAspThrIleIlePheArgSerThrAsp 378
Db 979 CTTCGTTGTTGCTTCTTTGAACTCTTGGTGGCCAGATGCTCAGCTGTTTCGGTCGACCGAC 1038
Qy 379 GlyGlyAlaThrThrArgIleTrpAspTrpThrSerTyrProAsnArgSerLeuArg 398
Db 1039 TCTGGGCAACATGAGCGGATCTGGCGGTGGCGAGCTATCCGACTGAGACCTATTATAC 1098
Qy 399 TyrValLeuAspIleSerAlaGluProTrpLeuThrPheGly---ValGlnProAsnPro 417
Db 1099 TACAGCATCTCACTCCCAAGACCGTGGATCAAGAACAACTTTATCGATGTGACGAGC 1158
Qy 418 ProValProSer-----ProLysLeuGlyTrpMetAspGluAlaMetAlaIle 433
Db 1159 GAGTCACCGTCCGATGCTCTCATCAGCGCTCGGCTGGATGATGATGATCTCTCAGATT 1218
Qy 434 AspProPheAsnSerAspArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsn 453
Db 1219 GACCAACCGCAGCAGCAACCACTGCTCTACGCGCACCGGAATGACAACTTTTGGCGGCCAC 1278
Qy 454 AspLeuThrLysTrpAspSerGlyGlnIleHisIleAlaProMetValLysGlyLeu 473
Db 1279 GATCTCAACATGGGACACGCGCCCAATGTGTCAATCCAACTCACTCGCAGACGCGCATC 1338
Qy 474 GluGluThrAlaValAsnAspLeuIleSerProSerGlyAlaProLeuIleSerAla 493
Db 1339 GAGGAATTCCTCGTCCAGGACCTGGCTCTGCACCCGCGGAGGAGGAGCTATTGGCCGCA 1398
Qy 494 LeuGlyAspLeuGlyGlyPheThrHisAla-----AspValThrAlaValProSer 510
Db 1399 GTCGAGCAGCAACACGCTTTCACCTTTGGCAGCAGCAACGACCTCGGAGACATCGCGCAG 1458

```

511 ThrIlePheThrSerProValPheThrThrGlyThrSerValAspTyrAlaGluLeuAen 530
 1459 ACGTCTGGGACGCCCATGCTGGCCACCTCGACAGCGTGCATAGCCGGGAACTCG 1518
 531 ProSerIleValAlaGlySerPheAspProSerSerGlnProAsnAspArgHis 550
 1519 GTCAAGAGCGTCTCGCGTGGCAACACCGCGGCACGACAG-----1563
 551 ValAlaPheSerThrAspGlyLysAsnTrpPheGlnGlySerGluProGlyGlyVal 570
 1564 GTGGCCATCTCTCGACGCGCGCGCGAGCTGAGCATCGATCGCGCGCGACAGTCC 1623
 571 ThrThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaProGly 590
 1624 ATGAACGGCGGACCGTGGCTTATTCGGCCGACGACACCATCTCTGTGTACCGCC 1683
 591 AspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGln 610
 1684 TCGTCCGGC-----GTGCAGCGCTCGCAGTTCCAGGCGAGCTTTCCTCGTCTCG 1734
 611 GlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThrPheTyrAla 630
 1735 AGCTTGGCGCGCGCGCGTCTATCGCTCGGACAAAGAACCAACAGCGTCTTTCACGCC 1794
 631 LeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThrPheGlnProValAla 650
 1795 GGCTTCGGATCGACCTTTTACGTCAAGAGGACACCGCGCAGCAGCTTC-----ACG 1845
 651 AlaGlyLeuProSerSerGlyAlaValGlyValMet-----PheHisAlaValProGly 668
 1846 CCGCGG-----CCCAAGCTGGCGAGCGGAGCGATTCGGGATATCGTCTCACCCGACC 1902
 669 LysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsnGlyGly 688
 1903 ACCGCGGCGACGTTGTATGTCTCGACCGAGCTCGCATATTCGCTCCACAGCTCGGCG 1962
 689 SerSerTrpSerAlaIle-----ThrGlyValSerSerAlaValAsnValGlyPheGlyLys 707
 1963 ACGACCTTTGGCCAAAGTCTCCACCGCCCTGACCAACACCTACACAGTCCGCTGGTGTG 2022
 708 SerAlaProGlySerSerTyrProAlaValPheValValGlyThrIleGlyValThr 727
 2023 GGCTCA-----GGCTCAACTGG---AACCTGTATGCTCTTCGGCACC-----GGCCCGTCA 2070
 728 GlyAla-----TyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAspAsp 745
 2071 GGGGCTCGCTCTACCGCAGTGGACACAGCGCGCTCTCTGACCGGACATCAGGGCTCC 2130
 746 GlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgArg 765
 2131 CAGGCTTCGGCTCCATCGACAGCACCAAGTTCGGCGGCGGCGGACCGCGCGGCA 2190
 766 ValTyrIleGlyThrAsnGlyArgGlyLeValTyr-----GlyAspIleGlyGlyAla 783
 2191 GTCTAGTGGGACCAACCGCGCGCGCTTTTACGCTCAGGGAACCGTCCGCGCGCGCG 2250
 784 ProSerGlySerProSerProSerValSerProSerAlaSerProSerLeuSerProSer 803
 2251 ACGGCGGGGACT-----TCTCGTCCACCAAGCAGAGC 2283
 804 ProSerProSerSerSerProSerProSerProSerProSerProSerSerSerSer 823
 2284 AGCAGCAGTACTCTCTCGCCAGCTCGAGCACCGCTGAGGTGAGGTGTTGATCCACG 2343
 824 ProSerProSerProSerProSerProSerProSerArgSerProSerProSerAlaSe 843
 2344 ACCCGGGCTTCGAGCGTGAATTCGTTCGAGGACACAGCTCGCGCGCGGTCACCGGGTCA 2403
 843 rProSerPro 846
 2404 GGGGTTCGCG 2413

RESULT 6

US-10-026-994-1
 ; Sequence 1, Application US/10026994
 ; Publication NO. US20030113732A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dunn-Coleman, Nigel
 ; APPLICANT: Goedegebuur, Frits
 ; APPLICANT: Ward, Michael
 ; APPLICANT: Yao, Jian
 ; TITLE OF INVENTION: EGY Endoglucanase and Nucleic Acids
 ; FILE OF INVENTION: Encoding the Same
 ; FILE REFERENCE: GC698
 ; CURRENT APPLICATION NUMBER: US/10/026,994
 ; CURRENT FILING DATE: 2002-04-30
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 2710
 ; TYPE: DNA
 ; ORGANISM: Trichoderma reesei
 US-10-026-994-1

Alignment Scores:
 Pred. No.: 8, 13e-105 Length: 2710
 Score: 1631.50 Matches: 353
 Percent Similarity: 57.2% Conservative: 130
 Best Local Similarity: 41.8% Mismatches: 297
 Query Match: 31.8% Indels: 64
 DB: 5 Gaps: 23

US-09-917-376-1 (1-957) x US-10-026-994-1 (1-2710)

QY 25 SerPheAlaValAlaAlaAlaLeuGlyValLeuProIleAlaIleThrAlaSerProAla 44
 Db 50 TCTCGAGCTTGTCCCTTGTCTCTGGGGCGGTC-----ATCCCTGCC 91
 QY 45 HisAlaAlaThrThrGlnProTyrThrTrpSerAsnValAlaIle---GlyGlyGlyGly 63
 Db 92 CATGCTGCC-----TTTTTCATGGAAGAACGTCAAGCTCGGCGCGCGCGCGC 139
 QY 64 PheValAspGlyIleValPheAsnGluGlyAlaProGlyIleLeuTyrValArgThrAsp 83
 Db 140 TTCGTCCTCCCGCATCATCTTCATCCCAAGCAAAAGGCGTAGCATATGCAGAACAGAT 199
 QY 84 IleGlyGlyMetTyrArgTrpAspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrp 103
 Db 200 ATTGGCGGCTGTACCGCTCAAC---GCCGACGACTCATGACCGCGCTCAGGATGG 256
 QY 104 Val-----GlyTrpAsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAla 119
 Db 257 ATTGCTGATAATGCGCGCTGGCACAACCTGG-----GGCATCGACGCTGTTCGCTT 307
 QY 120 AspProIleAsnThrAsnLysValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAsp 139
 Db 308 GATCCGAGGACGATCAAAAGGTGTATCGCGAGTCGCGCATGTATACGAAACAGCTGGAT 367
 QY 140 ProAsnAspGlyAlaIleLeuArgSerSerAspGlnGlyAlaThrTrpGlnIleThrPro 159
 Db 368 CCGAGTAATGAGCCATCATTCGCTCGTCAACCGCGCGGCAACGTTGCTTCACCAAC 427
 QY 160 LeuProPheLysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGlyArgLeuAlaVal 179
 Db 428 TTGCCCTTCAAGTCGGGGGTAAACATGACGACGCGGAGCGCGAGAGCGTCTGGCTGTC 487
 QY 180 AspProAsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArg 199
 Db 488 GATCCGGCAACTCCCAACATCATCTACTTGTGTCTGCTCAGGAAACGGCCTCTGGAAG 547
 QY 200 SerThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyr 219
 Db 548 TCTACGAGCGCGCGCTGACCTTTTCCCAAGGCTCTCGTCTTCACGGCAACTCGGACGTAC 607
 QY 220 IleAlaAsnProThrAspThrThrGlyTyrGlnSerAspIleGlnGlyValValTrpVal 239

608 ATCCAGACCCGAGTGATTCCAAACGGCTACACAGCGACAGCAAGGACTCATGTGGGTT 667
240 AlaPheAspLysSerSerSerLeuGlyGlnAlaSerThrIlePheValGlyVal 259
668 ACGTTGCACTCAACACGACGACGACCGGGGAGCCAGTCTCGTATCTTTTGGCAGC 727
260 AlaAspPro---AsnAsnProValPheThrPheArgAspGlyGlyAlaThrTrpGlnAla 278
728 GCTGATAACATCAGTCTTCACTGATGATGAGCAGCAATCGCGCTCCACGTGGAGTGT 787
279 ValProGlyAlaProThrGlyPheIleProHisGlyValPheAspProValAsnHis 298
788 GTACCGGGGACCCAGGGAATACCTTCTCACAAGGGAACATGACGACGAGAGAG 847
299 ValLeuThrIleAlaThrSerAsnThrGlyGlyProTyrAspGlySerSerGlyAspVal 318
848 GCCTTGATCTGACCTATTCGATGGCACAGGGCGGTATGATGCACATTTGGCTCAGTG 907
319 TrpLysPheSerValThrSerGlyThrTrpThrArgIleSerProValProSerThrAsp 338
908 TGGAGGTACGACATTCGACAGGGGGAATTTGAAAGACATCACCCCTGTCTCTGGATCAGAT 967
339 ThrAlaAsnAspTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThr 358
968 CTA-----TACTTTGGCTTTGGCGGCTTGGCTTCGATTGCAAAAGCCAGGAACC 1018
359 IleMetValAlaThrGlnIleSerTrpTrpProAspThrIleIlePheArgSerThrAsp 378
1019 CTTGTGTGTCTTTTGAACCTCTTGGTGGCAGATGCTCAGCTGTTCTGGTTCAGCCGAC 1078
379 GlyGlyAlaThrTrpThrArgIleTrpAspTrpThrSerTyrProAsnArgSerLeuArg 398
1079 TCTGGGACAAATGAGCGCCGATCTGGCGGTGGCGAGCTATCCGACTGACACCTATTATC 1138
399 TyrValLeuAspIleSerAlaGluProTrpLeuThrPheGly---ValGlnProAsnPro 417
1139 TACAGCATCTCAACTCCCAAGACCGGTGATCAAGAACAACTTTATCGATGTGACGAGC 1198
418 ProValProSer-----ProLysLeuGlyTrpMetAspGluAlaMetAlaIle 433
1199 GAGTCACGTCGATGTTCTCATCAGCGCTCTGGCTGGATGATGATGATCTCTCGAGATT 1258
434 AspProPheAsnSerAspArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsn 453
1259 GACCAACCGACAGCAACCACTGCTCTACGGCACCGGAATGACAATCTTTTGGCGGCAC 1318
454 AspLeuThrLysTrpAspSerGlyGlyGlnIleHisIleAlaProMetValLysGlyLeu 473
1319 GATCTCAACCACTGGACACGCGCCCAATGTGTCAATCCCAATCACTGGCAGACGGCATC 1378
474 GluGluThrAlaValAsnAspLeuIleSerProSerGlyAlaProLeuIleSerAla 493
1379 GAGGAATCTCCGTCAGAGCTGGCTCTCCACCGCGGAGGAGCTATTTGGCGGCA 1438
494 LeuGlyAspLeuGlyGlyPheThrHisAla-----AspValThrAlaValProSer 510
1439 GTCGAGACGACACACGGCTTCACTTTGCCAGCAGAAACGACCTCGGGACATCCCGCAG 1498
511 ThrIlePheThrSerProValPheThrThrGlyThrSerValAspTyrAlaGluLeuAsn 530
1499 ACGGTCTGGGCAACGCCACATCGGCGCCACCTCGACGCGTCTGACTACGCGGGAATCG 1558
531 ProSerIleIleValArgAlaGlySerPheAspProSerSerGlnProAsnAspArgHis 550
1559 GTCAGAGCGTGTCCGCTCGGCAACACCCCGCGCACGCAACAG----- 1603
551 ValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGlyVal 570
1604 GTGGCCATCTCTCGACGGCGCGCGGACGCTGGAGCATCGACTACGCGCGCGCACAGCTCC 1663
571 ThrThrGlyGlyThrValAlaAlaSerAlaAlaAspGlySerArgPheValTrpAlaProGly 590
1664 ATGAACGGCGCACGGTGGCTATTTCGGCGGACGCGCACGATCTCTGTGTCGACCGCC 1723

591 AspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGln 610
1724 TCGTCCGCGC-----GTGACGGCTCGGAGTTCCAGGCGAGCTTTGCTCCCTCTCG 1774
611 GlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThrPheTyrAla 630
1775 AGCTCCCGCGCGCGCTCATCGCTCGAACAAGACCAACAGCGCTCTTCTACGCGC 1834
631 LeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThrPheGlnProValAla 650
1835 GGCTCCGGATCGACTTTTACGTACGAAGACACCGCAGCAGCTTC-----ACG 1885
651 AlaGlyLeuProSerSerSerGlyAlaValGlyValMet-----PheHisAlaValProGly 668
1886 CGCGCG---CCCAAGCTGGCGCAGCGCAGGACGATCGCGGATATCGCTGCTCACCCGACC 1942
669 LysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsnGlyGly 688
1943 ACCCGGGCAGCTTGTATGTCTCGACCGACGCTCGGCATATTCGCTCCACAGACTCGGCG 2002
689 SerSerTrpSerAlaIle---ThrGlyValSerSerAlaValAsnValGlyPheGlyLys 707
2003 ACGACCTTTGGCCAAAGTCTCCACCGCTGACCAACACTACAGATCGCCCTGGGTGTG 2062
708 SerAlaProGlySerSerTyrProAlaValPheValValGlyThrIleGlyValThr 727
2063 GGCTCA---GACTCGAAGTGG---AAGCTGTATGCTTCGGCACC-----GGCCCGTCA 2110
728 GlyAla-----TyrArgSerAspAspCysGlyThrTrpValLeuIleAsnAspAsp 745
2111 GGGGCTCGCTCTACCGCAGTGGAGACAGCGCGCTCTCGACGCGACATCCAGGGCTCC 2170
746 GlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArg 765
2171 CAGGGCTTCGGCTCCATCGACAGCACCAAGTTCGCGCAGCGCAGCACCGCGGGCAA 2230
766 ValTyrIleGlyThrAsnGlyValGlyIleValTyr-----GlyAspIleGlyGlyAla 783
2231 GTCTACGTGGCACCAACCGCGGGGCTCTTTTACGCTCAGGGAACCGTTCGGCGCGCGC 2290
784 ProSerGlySerProSerProSerValSerProSerAlaSerProSerLeuSerProSer 803
2291 ACGGGCGGACT-----TCTCGTTCGACCAACGACAGCAGC 2323
804 ProSerProSerSerProSerProSerProSerProSerProSerProSerSerSer 823
2324 AGCAGCAGTACTCTTCGCGCAGCTCGAGCACACGCTGAGGTGAGCGTTGTATCCACG 2383
824 ProSerProSerProSerProSerProSerProSerProSerProSerProSerAlaSe 843
2384 ACCCGGCTTCGAGCGTGACTTCGTGAGGACAGCTCGCGCGCGCTCCACCGGGGTCA 2443
843 rProSerPro 846
2444 GGGGTCCGCG 2453

RESULT 7

US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 1
 ; LENGTH: 9025608
 ; TYPE: DNA
 ; ORGANISM: Streptomyces avermitilis
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (4187715)
 ; OTHER INFORMATION: a, t, c, g, other or unknown
 US-10-156-761-1

Alignment Scores:

Pred. No.: 8,06e-97 Length: 9025608
 Score: 1570.00 Matches: 393
 Percent Similarity: 49.2% Conservative: 131
 Best Local Similarity: 36.9% Mismatches: 346
 Query Match: 30.6% Indels: 194
 DB: 6 Gaps: 38

US-09-917-376-1 (1-957) x US-10-156-761-1 (1-9025608)

QY 4 SerGluAsnIleArgLeuThrMetArg-----SerArgArgLeuValSerLeu 19
 DB 3162026 TCTCCAGAACGAGGTATGTATCGCACGCGCCCGCGGACGACGAAACCGTC----- 3161973
 QY 20 LeuAlaIleThrAlaSerPheAlaValAlaAlaLeuGlyValLeuProIleAlaIle 39
 DB 3161972 CTCGCGGGACCGCGCG-----GCCCGCGGCTTCACGCGCGTCCCG-----GCC 3161928
 QY 40 ThrAlaSerProAlaHisAlaAlaThrThr-----GlnProTyrThrThrSerAsn 56
 DB 3161927 GTGGGGTCCGAGGCGACCGCGCGCAACACCGCGCGCGCTGTACCGTGGCGCAAC 3161868
 QY 57 ValAlaIleGlyGlyGlyPheValAspGlyIleValPheAsnGluGlyAlaProGly 76
 DB 3161867 GCCGTATCGGGGACCGGCTTCGTACCGCGGTCTCTCCACCCCTCGGTACGCGGT 3161808
 QY 77 IleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTTPAspAlaAlaAsnGlyArg 96
 DB 3161807 CTCGCTACGCGGACCGACATCGGCGGCTTACCGTGGGACGACCGCGCGCGC 3161748
 QY 97 TrpIleProLeuLeuAspTTPValGlyTrpAsnAsnTTPGlyTyrAsnGlyValValSer 116
 DB 3161747 TGGACCGGCTCATCGACCACTCGGTGGGACGACTGGAACTCTCTCGCGTTCGAGCG 3161688
 QY 117 IleAlaAlaAspProIleAsnThrAsnLysValTTPAlaAlaValGlyMetTyrThrAsn 136
 DB 3161687 ATGGCGGTGACCCCGACCGACCGCGGTCTACCTCGCGTGGGACCTACGCGCG 3161628
 QY 137 SerTTPAspProAsnAspGlyAlaIleLeuArgSerSerAspGlnGlyAlaThrTTPGln 156
 DB 3161627 TCGTGGGCGGCAAC---GGCGCGTCTCGCTCCGAGGACCGCGCGCGCACCTGGAC 3161571
 QY 157 IleThrProLeuProPheLeuGlyGlyAsnMetProGlyArgGlyMetGlyGluArg 176
 DB 3161570 CGCACCGACTGACCGTGAAGCTCGGCGGCAACGAGGACGCGCGCGCGCGTTCGAGCG 3161511
 QY 177 LeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGly 196
 DB 3161510 CTCCTGTGACCGCGCGGACGACACCTCTGGTGGGCGACG---CGGACGACGCGG 3161454
 QY 197 LeuTTPArgSerThrAspSerGlyAlaThrTTPSerGlnMetThrAsnPheProAspVal 216
 DB 3161453 CTGCTCAAGTCGACCGACCGGGGCGCCACTTGGGACGCGCGCGCGCTTCCG----- 3161400
 QY 217 GlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGlnSerAspIleGlyVal 236
 DB 3161399 -----CGGAAGGCGAACTCTCTCGCGGCGAGGAGTC 3161370

QY 237 ValTTPValAlaPheAspLysSerSerSerSerLeuGlyGlnAlaSerLysThrIlePhe 256
 DB 3161369 GTGTTC-----CTCGTCGCCCGCGCGACCGTCTAC 3161337
 QY 257 ValGlyValAlaAspProAsnAsnPro-----ValPheTTPSerArgAsp 271
 DB 3161336 GCCGCTGGGTGACGGCGACCGCACCTCGGGCAGCGGCAACCTGTACCGCACGCGCAC 3161277
 QY 272 GlyGlyAlaThrTTPGlnAlaValProGlyAlaProThrGlyPhe-----IlePro 288
 DB 3161276 ---GGCACGACCTGGGGCGCGTCCCGCGCGCGCTCCGGCACCTCCGGCAAGGTCCCG 3161220
 QY 289 HisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsnThrGly 308
 DB 3161219 CTCGCGCGCGGTACGACACGACACCGCGGAGCTGTAGCTACGCGACGCGCACCC 3161160
 QY 309 GlyProTyrAspGlySerSerGlyAspValTTPLysPheSerValThrSerGlyThrTTP 328
 DB 3161159 GGCCCGGGCGGCGAGTTCGACGCGGAGCGTGCACAGCTCGGTACCGCCACGCGGACGTGG 3161100
 QY 329 ThrArgIleSerProVal-----ProSerThrAspThrAlaAsnAspTyrPhe 344
 DB 3161099 ACCGAGGTCACTCCCGGTGAAGCGGGCGGACGACGACGCGCTCGCGCGACCTTC 3161040
 QY 345 GlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGln 364
 DB 3161039 GCCTACGCGGGGTCCGCTGACGCGCGCGCGCGCGCGCTCGCTCGTCTCCACCAAC 3160980
 QY 365 IleSerTTPProAspThrIleIlePheArgSerThrAspGlyGlyAlaThrTTPThr 384
 DB 3160979 AACCGCTGGCGGCGGCGGACGCGTCTTCGCTCCACGCGCGCGCGCTACTCGGACG 3160920
 QY 385 ArgIleTTPAspTTPThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSer 404
 DB 3160919 TCCCTCAAGGAC-----GCCCGCGTGTTCGACGTGTC 3160887
 QY 405 AlaGluProTTPLeuThrPheGlyValGlnProAsnProProValProSerProLysLeu 424
 DB 3160886 GAGACTCCCTTCTCGACTGGGCGGCGACGAC-----AAGCGGAAGTTC 3160845
 QY 425 GlyTTPMetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeuTyrGly 444
 DB 3160844 GGTGTGTGATTCAGCGCGTTCGCGTTCGACCGTACGACTCCCGACGACGTCTGTACGGG 3160785
 QY 445 ThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTyrAspSerGlyGlyGlnIle 464
 DB 3160784 ACCGCGCGGACCTCTACGCGACCGCGACCTCAAGCGCTGG----- 3160743
 QY 465 HisIleAlaProMetValLysGlyLeuGluThrAlaValAsnAspLeuIleSerPro 484
 DB 3160742 -----GCACCGCGGATCGCGGCTTGAGGAGAGCGCGCTGCCCACTGATCTGCC 3160689
 QY 485 ProSerGly---AlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAla 503
 DB 3160688 CCGGTGGGAGGACACCTGTATCAGCGACTCGGGGACATCGGTGTGTGTACCGAG 3160629
 QY 504 AspValThrAlaValProSerThr---IlePheThrSerProValPheThrThrGlyThr 522
 DB 3160628 CGGTTCAGCGGCTCTCCGTTCGCGCGGATGGCGAGCAACCCCGTGTTCGGGTCGGGACG 3160569
 QY 523 SerValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspPro 542
 DB 3160568 GGACTCGCGGAGCGCGCGCGCGCGGTATGTCTCGCACGCGGCTGGGCGGCGACCC 3160509
 QY 543 SerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTTP--- 561
 DB 3160508 GGCAACGCG-----GGGTACTCCACGACGCGCGCGGCGGACCTGGGCG 3160467
 QY 562 ---PheGlnGlySerGluProGlyValThrThrGlyThrGlyThrValAlaAlaSerAla 580
 DB 3160466 CCCTTCGAGGCGCGCGCGGACATCGCAAGGACGACCGCGGACCGATCGCCACGATGCC 3160407
 QY 581 AspGlySerArgPheValTTPAla-----ProGlyAspProGlyGlnProVal 596

```
Db 3160406 GACGGGGGACACTGCTGTGGTCTTCTGTCACCTGGAGCGGC-----ACG 3160362
Qy 597 ValTyrAla-----ValGlyPheGlyAsnSerTyrAlaAlaSerGlnGlyVal 612
Db 3160361 ACGTAGCGCGCCACCGCTCGACGAGCAACGGCGGAGCTGGTCCGAGGTCTCTCTCTTC 3160302
Qy 613 ProAlaAsnAlaGlnIleArgSerAspArgValAsnProIysThrPheTyrAla----- 630
Db 3160301 CCGAAGGGCGCCACACCGGTGCGCGACCCCGCGCATCCGACGGCTTCTACGCGCTACGAC 3160242
Qy 631 LeuSerAsnGlyThrPheTyrArgSerThrAspGlyValThrPheGlnProValAla 650
Db 3160241 TTCGACAAATGGAACTATACCGCAGCACTGACAGTGGCGGTTCGTTCACGGCCCGTGGC 3160182
Qy 651 AlaGlyLeuProSerSerGlyAlaValGlyValMetPheHisAlaValProGlyLysGlu 670
Db 3160181 GCGGAGCTGCC-----TCCGGGACACCGACTTCAAGCTGTGCGCGCGCGGACGAGC 3160125
Qy 671 GlyAspLeuTyrPheAlaAlaSer---SerGlyLeuTyrHisSerThrAsnGlyGlySer 689
Db 3160124 GCGACCTGTGGCTCTCGCGCAATGGAACGGGCTCTACCGGTCCACCGACGGCGGGAC 3160065
Qy 690 SerTyrSerAlaIleThrGlyValSerSerAlaValAsnValGlyPheGlyLysSerAla 709
Db 3160064 ACCTTGCGCGAGGATCGACAGCTGCTGGCTCTACACCTCGGCTTCGCAAGGGCGCC 3160005
Qy 710 ProGlySerSerTyrProAlaValPheValValGlyThrIleGlyGlyValThrGlyAla 729
Db 3160004 GACGGCGCGACATACCGCGGATCTACAGGTGCGTCCAGGAGACCATCACCGCGCTC 3159945
Qy 730 TyrArgSerAspAspCysGlyThrTyrThrTyrValIleAsnAspAspGlnHisGlnTyr 749
Db 3159944 TACCGCTCCGACGACGCCCGCAGACATGGTCCGGATCAACGACGCGCCACCATGG 3159885
Qy 750 GlyAsnTyrGlyClnAlaIleThrGlyAspHisAlaAsnLeuArgValTyrIleGly 769
Db 3159884 GGGTGTATCGCGGAGCGCGTCTGTCGTGACCCCGCATCCACGGCGGGTCTACCTCGCC 3159825
Qy 770 ThrAsnGlyArgGlyIleValTyrGlyAspIle-----GlyGlyAlaPro 784
Db 3159824 ACCAAGCGCGCGGATCAGTACGGGGAGCGGCTGTGATCCCGCGGCTCACGGATGCCA 3159765
Qy 785 SerGlySerProSerProSerValSerProSerAlaSerProSerLeuSerProSerPro 804
Db 3159764 CCGGGCGCGCATCTGTTCGTGGCG---ACTACACCCCGGACGAGTGGCCCGAGGAG 3159708
Qy 805 Ser-----ProSerSerProSerPro----- 812
Db 3159707 TCTGGCAGGAGGACGTACGCCTCATGAAGGAGCGCGCGTCAACTCCGTACCGTCCGGC 3159648
Qy 813 ---SerProSerProSerSerPro----- 820
Db 3159647 TCTTCTCTGGGCGGAGCTCGAACCCACCCCGGGGCGAGGAGTTCGGGTGCTCGACC 3159588
Qy 821 -----SerSerSerProSerProSerPro 828
Db 3159587 GCCTATGACACGTATGACAGAGAACGGCATCGCGCTGCTGTCGGCCACCCCGCCCT 3159528
Qy 829 SerProSerPro----- 832
Db 3159527 CGCCCCCGCTGGATGGGGCGCTGCACCCCGAGACCTGCCCGCGAGGAGCGGCC 3159468
Qy 833 -----SerProSerArgSerProSerProSerAla 842
Db 3159467 GCATCGAGTGTGGGCGCGCGCAGACATTCGCCCACTCCAGTGGCGGTCTACCGCGCT 3159408
Qy 843 SerProSerProSerSerProSerProSerProSerSer-----ProSerSer 858
Db 3159407 ACGCCGCGCGCCATCATCCGAGGACCTCGCGCGCGCTACCGCGCGCCATCCCGCTCACGA 3159348
Qy 859 SerProSerProThrProSerSerProValSerGlyGlyValLys----- 874
```

```
Db 3159347 TGTGGCACATCAACACAGTACTGACCTACGACTGGGGCGACAGGGCGGCCACCGCCT 3159288
Qy 875 -----ValGlnTyrLysAsnAsnAspSerAlaProGlyAsp 886
Db 3159287 TCCGCCGCTGGCTCCAGGGCAGGTACGCGACGCTCGACGACTCAACACGGCTGGG--- 3159231
Qy 887 AsnGlnIleLysPro---GlyLeuGlnValValAsnThrGlySerSerValAspLeu 905
Db 3159230 -----GCACGGCTTCTGGAGCCAGGCTACGACGACTGGGAGGACATCTCCCGCCGC 3159177
Qy 906 SerThrValThrValAlaGlyTyrTyrPheThrArgAspGlyGlySerSerThrLeuValTyr 925
Db 3159176 GCCACGCCCACTACATGAACACCCACCCAGTCTCGACTTCAAGCGCTTCACTCCG 3159117
Qy 926 AsnCysAspTyrAlaAlaIleGlyCysGlyAsnIleArgAlaSerPheGlySerValAsn 945
Db 3159116 ACGCACTCATGGAGTGT---TCGTGCCGAGACGTGACATCTGCTG--- 3159075
Qy 946 ProAlaThrPro 949
Db 3159074 CCGGCACACCC 3159063

RESULT 8
US-10-156-761-2561
; Sequence 2561, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 2561
; LENGTH: 2217
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2217)
US-10-156-761-2561

Alignment Scores:
Pred. No.: 1.33e-91 Length: 2217
Score: 1442.00 Matches: 328
Percent Similarity: 54.2% Conservative: 107
Best Local Similarity: 40.8% Mismatches: 272
Query Match: 28.1% Indels: 96
DB: 6 Gaps: 25

US-09-917-376-1 (1-957) x US-10-156-761-2561 (1-2217)
Qy 7 IleArgLeuThrMetArgSerArgLeuValSerLeuLeuAlaAlaThrAlaSerPhe 26
Db 1 ATCGCACGCGCGCGCGGAGCAGACGAAACCGTC-----CTCGCGGAGACCGCGCG--- 51
Qy 27 AlaValAlaAlaLeuGlyValLeuProIleAlaIleThrAlaSerProAlaHisAla 46
Db 52 -----GCCGCGCGCTCACGCGCTCCCC-----GCCGTGGGTGCCAGCGCACGCC 99
Qy 47 AlaThrThr-----GlnProTyrThrTyrSerAsnValAlaIleGlyGlyGly 63
Db 100 GCCGAACACCGCGCGCGCTCGTACGCTGGCGAACACCGCGCTCATCGGGGCGCACCGGC 159
```

QY 64 PheValAspGlyIleValPheAsnGluGlyAlaProGlyIleLeuTyrValArgThrAsp 83
DB 160 TTCGTACCGCGGTCTCTCCACCCCTCCGTACCGGTCTCGCTACCGCGGACCGAC 219
QY 84 IleGlyGlyMetTyrArgTyrAspAlaAlaAsnGlyArgTyrIleProLeuLeuAspTyr 103
DB 220 ATCCGCGGCGCTACCGCTGGACACCGCGGCGCCCTGGACCCCGCTATCAGCCAC 279
QY 104 ValGlyTyrAsnAsnTyrGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsn 123
DB 280 CTCGGCTGGACGACTGAACTCTCCGCGTGGAGCGATGGCGCTCGACCCACCGAC 339
QY 124 ThrAsnLysValTyrAlaAlaValGlyMetTyrThrAsnSerTyrAspProAsnAspGly 143
DB 340 CCGGACCGGCTTACCTCCGCGTGGCACCTACGCCAGTCTGTGGCGGGCAAC---GGC 396
QY 144 AlaIleLeuArgSerAspGlnGlyAlaThrTyrGlnIleThrProLeuProPheLys 163
DB 397 GCGGTCTGCGCTCCGAGGACCGCGCGCCACCTGGACCCGACCTGACCGTGAAG 456
QY 164 LeuGlyGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsn 183
DB 457 CTCGGCGGCAACGAGGACCGCGCGCGCGGTGAGCGACTCTCTGTCGACCCCGCGCGAC 516
QY 184 AspAsnIleLeuTyrPheGlyAlaProSerGlyLysGlyLeuTyrArgSerThrAspSer 203
DB 517 AGCGACACCTCTGGCTGGCGACG---CGGACGACGCGGCTGCTCAAGTCGACCGACCGG 573
QY 204 GlyAlaThrTyrSerGlnMetThrAsnPheProAspValGlyTyrIleAlaAsnPro 223
DB 574 GCGCGCACTTCGGCGACCGCGACCGCTTCCCG----- 606
QY 224 ThrAspThrThrGlyTyrGlnSerAspIleGlnGlyValValTyrValAlaPheAspLys 243
DB 607 -----GCGAGCGCACTCTCCGGGAGGAGTCGTGTC----- 642
QY 244 SerSerSerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspProAsn 263
DB 643 -----CTCGTCGCGCGCGCGCGCACCGCTACCGCGGTGGGGTACCGCGCGAC 690
QY 264 AsnPro-----ValPheTyrSerArgAspGlyGlyAlaThrTyrGlnAla 278
DB 691 GGCACCTCGCGGACCGCGCACTGTACCGCGACCGCGGAC---GGCACGACCTGGGGGCGC 747
QY 279 ValProGlyAlaProThrGlyPhe-----IleProHisLysGlyValPheAspPro 295
DB 748 GTCCCGCGCGCGCTCCGCGACCTCCGCCAAGTTCCTCCGCGCGCGGTACGACACG 807
QY 296 ValAsnHisValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyrAspGlySer 315
DB 808 CACACCGCGGAGCTGTAGTACGTACGGGACGACCGCGCGCGCGCGCGCGCGTCCGAC 867
QY 316 GlyAspValTyrLysPheSerValThrSerGlyThrTyrArgIleSerProVal--- 334
DB 868 GGCAGCGTGCAAGTCTCGTACCGCGACCGCGGCGTGGACCGAGGTACACCGCGTGAAG 927
QY 335 -----ProSerThrAspThrAlaAsnAspTyrPheGlyTyrSerGlyLeuThrIle 351
DB 928 CCGGCGGGGACGACGAGCGGCTCGCGCGACACCTTCGCTACGCGGGGTCCCGCTC 987
QY 352 AspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTyrTyrProAspThr 371
DB 988 GACGCGCGCGCGCGCGCGCGCTGTCTCTCCACCAACACCGCTGGCGCGCGCGCGAC 1047
QY 372 IleIlePheArgSerThrAspGlyAlaThrTyrThrArgIleTyrAspTyrThrSer 391
DB 1048 ACGGTCTCCGTTCCGTTCCAGGACCGCGCGCGGTACCTGGAGTCTCCCTCAAGAC----- 1098
QY 392 TyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGluProTyrLeuThrPhe 411
DB 1099 -----GCGCGCGTGTTCGACGTGTCCGAGCTCCCTCTCTCGACTGG 1140

QY 412 GlyValGlnProAsnProValProSerProLysLeuGlyTyrMetAspGluAlaMet 431
DB 1141 GCGGACGAC-----AAGCCGAAGTTCGGCTGGTGGATCCAGCGCGCTC 1182
QY 432 AlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAla 451
DB 1183 GCGGTGACCGCTACGACTCCAGCAGCTGCTGTACGGGACCGCGCGGACCTCTACGGC 1242
QY 452 ThrAsnAspLeuThrLysTyrAspSerGlyGlyGlnIleHisIleAlaProMetValLys 471
DB 1243 ACCCGCGACTCAAGCGCTGG-----GCACCGCGGATCCGC 1278
QY 472 GlyLeuGluGluThrAlaValAsnAspLeuIleSerProProSerGly---AlaProLeu 490
DB 1279 GCGCTGGAGGAGCGCGCTGCGCAACTGATCTCGCCCGCGCGGGAGGACACCTG 1338
QY 491 IleSerAlaLeuGlyAspLeuGlyPheThrHisAlaAspValThrAlaValProSer 510
DB 1339 ATCAGCGGACTCGGGACATCGGTGTATGTACACGCGGCTCAGGGCTCTCCGTCG 1398
QY 511 Thr---IlePheThrSerProValPheThrThrGlyThrSerValAspTyrAlaGluLeu 529
DB 1399 CCGCGCATCGCGACGAAACCCGCTGTCGGTTCGGCGCGGACTCGCGCAGCGCGCGCC 1458
QY 530 AsnProSerIleLeuValArgAlaGlySerPheAspProSerSerGlnProAsnAspArg 549
DB 1459 AGCGCGCGCTATGTCGTCCGCGCGGCTGGGCGCACCGCGCAACGCGC----- 1506
QY 550 HisValAlaPheSerThrAspGlyLysAsnTyr-----PheGlnGlySerGluPro 567
DB 1507 -----GCGTACTCCACGCGCGCGGCGACTTGGCGCGCTTCGAGGCGCGACCGCGAC 1560
QY 568 GlyGlyValThrThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheValTyr 587
DB 1561 ATCGCAAGGACGCGCGGACCGATCGCCACCGAGTCCGCGCGCGCGCACCTGCTGTGG 1620
QY 588 Ala-----ProGlyAspProGlyGlnProValValTyrAla----- 599
DB 1621 TCCTTCGTGCACTGGGACGCG-----ACGACGTACGCGGCGCGCGCGCTCG 1665
QY 600 ValGlyPheGlyAsnSerTyrAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArg 619
DB 1666 ACGGCAACCGCGCGGAGTGTCTCCGAGTCTCTCTCCCGAAGGCGCGCACCGCGTC 1725
QY 620 SerAspArgValAsnProLysThrPheTyrAla-----LeuSerAsnGlyThrPheTyr 637
DB 1726 GCGGACCGCGCGGCTTCTACGCGCTTACGACTTCGACATGGAAGCGTATAC 1785
QY 638 ArgSerThrAspGlyGlyValThrPheGlnProValAlaAlaGlyLeuProSerSerGly 657
DB 1786 GCGAGCACTGACAGTGGCGCTTCGTTACGCGCGCGCGCGCGGCGGACTGCCC---TCGCGC 1842
QY 658 AlaValGlyValMetPheHisAlaValProGlyLysGluGlyAspLeuTyrLeuAlaAla 677
DB 1843 GACAGCCAGTTCAGAGTGTGTCGCGCGCGCGCGGACGAAAGCGCGCTGTGGCTCTCCGCGC 1902
QY 678 Ser---SerGlyLeuTyrHisSerThrAsnGlyGlySerSerTyrSerAlaIleThrGly 696
DB 1903 AAATGGAACGGGCTCTACCGGTCCACCGCGCGCGGACACCTTCGCGAGGATCGACAGC 1962
QY 697 ValSerSerAlaValAsnValGlyPheGlyLysSerAlaProGlySerSerTyrProAla 716
DB 1963 TCCTGGGCTCTGTACACCTCGCTTCGCGAAGGCGCGCGCGCGCGCGGCGGCGG 2022
QY 717 ValPheValValGlyThrIleGlyValThrGlyAlaTyrArgSerAspAspCysGly 736
DB 2023 ATCTACCGAGTTCGCTCGACGCGGACCATCACCGCGCTTACCGCTCCGACGCGCGCC 2082
QY 737 ThrThrTyrValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnTyrGlyGlnAlaIle 756
DB 2083 AGGATGCGGTCCGGATCAACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 2142
QY 757 ThrGlyAspHisAlaAsnLeuArgValTyrIleGlyThrAsnGlyArgGlyIleVal 776

```
Db 2143 GTGGGTGACCCGCGCATCCAGCGCGGTCTACTCCCAACAGCGCGCGCATCCAG 2202
Qy 777 TyrGlyAsp 779
Db 2203 TAGCGGGAG 2211

RESULT 9
US-09-927-827-21
; Sequence 21, Application US/09927827
; Publication No. US20030036176A1
; GENERAL INFORMATION:
; APPLICANT: Bower, Stanley G.
; APPLICANT: Ramseier, Thomas M.
; TITLE OF INVENTION: Directed Genetic Engineering of Xanthomonas campestris
; FILE REFERENCE: 38-10(15824)B
; CURRENT APPLICATION NUMBER: US/09/927,827
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/279,493
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 69
; SEQ ID NO 21
; LENGTH: 3668
; TYPE: DNA
; ORGANISM: Xanthomonas campestris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1001)..(2668)
US-09-927-827-21

Alignment Scores:
Pred. No.: 4,18e-72 Length: 3668
Score: 1167.50 Matches: 299
Percent Similarity: 48.6% Conservative: 137
Best Local Similarity: 33.3% Mismatches: 351
Query Match: 22.7% Indels: 112
DB: 3 Gaps: 28

US-09-917-376-1 (1-957) x US-09-927-827-21 (1-3668)
Qy 23 ThrAlaSerPheAlaValAlaAlaLeuGlyValLeuProIleAlaIleThrAlaSer 42
Db 1136 ACCACAACAGGAGCATGGTGGTGTCTTACGCGCTGCTG---CTGTGCTCTTTTCGACG 1192
Qy 43 Pro-----AlaHisAlaAlaThrGlnProTyrThrTrpSerAsnValAla 58
Db 1193 CCAGCGTGGTCCGCGGAGCCCGCCACGTCGCGGCGCTTACCAGTGGCGAGGTGCGCC 1252
Qy 59 IleGlyGlyGlyPheValAspGlyIleValPheAsnGluGlyAlaProGlyIleLeu 78
Db 1253 ATTGGCGGTGGCGCTTTGTACCGGTGTGCTGTTTCATCCCGCGCAACGTGTCTGGCC 1312
Qy 79 TyrValArgThrAspIleGlyMetTyrArgTrpAspAlaAlaAsnGlyArgTrpIle 98
Db 1313 TATCGCGCACCATGTGGTGGCGGTACCGCTGGGATCGCGAGCGCGAGCATGGACC 1372
Qy 99 ProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyrAsnGlyValValSerIleAla 118
Db 1373 GCGCTGACGACTGGTGGCGGTGTACGACTGGAACCTGTATGGGCAATCGCATTCGCC 1432
Qy 119 AlaAspProIleAsnThrAsnLysValTrpAlaAlaValGlyMetTyrThrAsnSerTrp 138
Db 1433 CTCACCCCGCGCATGCCGATGGCTGTATCTGGCGCGCGCACCTATATGCATGAA--- 1489
Qy 139 AspProAsnAspGlyAlaIleLeuArgSerSerAspGlnGlyAlaThrTrpGlnIleThr 158
Db 1490 CGCGCGCGCACTGCCGAGTTGGCTCTGTCAACCGCGCGCGCAGTTCGAGCGTGC 1549
Qy 159 ProLeuProPheLysLeuGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAla 178
Db 1550 GACCTCCCGTTTAAGCTGGGTGTAAACCACTGGCGCGCGCAATGATGGCGAGCGCTGGCG 1609
Qy 179 ValAspProAsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGlyLeuTrp 198
```

```
Db 1610 GTGACCCGCGCATGGCGCGGTGCTGCTGCTGGGCTCGCGGATGCC---GGCCTGTGG 1666
Qy 199 ArgSerThrAspSerGlyValAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThr 218
Db 1667 CGTAGCGAGCATCGCGCGCGCATCGGGCGAGGTGGCGTCTTTCCGAGCCGCGCGTG 1726
Qy 219 TyrIleAlaAsnProThrAspThrThrGlyTyrGlnSerAspIleGlnGlyValValTrp 238
Db 1727 GCCGTGTCACCGCGCGCAATCATGTTGGCGCGAGCAGCGGTG---GGGATCGCCTTT 1783
Qy 239 ValAlaPheAspLysSerSerSerSerLeuGlnAlaSerLysThrIlePheValGly 258
Db 1784 GTCGTGTTTCGAGCGAGCATGGCAACACGCGTCCCAACACCGCGCATCTACGTGGGC 1843
Qy 259 ValAlaAspProAsnAsnProValPheTrpSerArgAspGlyGlyAlaThrTrpGlnAla 278
Db 1844 GTGTCCACCGAACACAGCAGCGCTGTATGTGTCGGAAGATGCGCGCGCATGTTGGCACCG 1903
Qy 279 ValProGlyAlaProThrGlyPheIleProHisLysGlyValPheAspProValAsnHis 298
Db 1904 GTGCGCGGCAACACCGCGGCTGCGCCCGAGCCACATGCGCGCGGAGCATGGGCAC 1963
Qy 299 ValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyrAspGlySerSerGlyAspVal 318
Db 1964 ---TGTATCTGAGTATGGCGACCGAGCCGCGCGGACCTGATGCGCGGGGAGCGCTTG 2020
Qy 319 TrpLysPheSerValThrSerGlyThrTrpThrArgIleSerProValProSerThrAsp 338
Db 2021 TGSAAATTCACGCGCGCACAGGCGCGTGGCGTGAGATCAGCCCGATTCGCGAG---CCA 2077
Qy 339 ThrAlaAsnAspTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThr 358
Db 2078 GCCAGTGGCGATGATTCGCTGGGTGGGTGGCGGTGGATCCGCAACATCCGAGGTG 2137
Qy 359 IleMetValAlaThrGlnIleSerTrpTrpProAspThrIleIlePheArgSerThrAsp 378
Db 2138 CTGCTGGCGCACCATCTTCGCGCGTCGCGAGCGCGCGAGCTGTATCGAGCGGTGAT 2197
Qy 379 GlyGlyAlaThrTrpThrArgIleTrpAspTrpThrSerTyrProAsnArgSerLeuArg 398
Db 2198 GGTGGCAAGCATCGAGCGCGCTG---TTGGCCGAT 2230
Qy 399 TyrValLeuAspIleSerAlaGluProTrpLeuThrPheGlyValGlnProAsnProPro 418
Db 2231 CGCGTGTTCGATCACAGCGCGCGCGCGG---ACCGCACATGCCCGCGCAC---2281
Qy 419 ValProSerProLysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSer 438
Db 2282 -----TGGATGGG---GGCTGGCGCATCGATCCGTTCGACGCG 2317
Qy 439 AspArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrp 458
Db 2318 AACCATGCGCTGTTCGTACCGCGTACCGCATCTGGGCTCGGCGCATCTG---2368
Qy 459 AspSerGlyGlyGlnIleHisIleAlaPro-----MetVal 470
Db 2369 -----CAGGATTCGCGCACCGCGCGCGCTGCACTGTGGTGTTCAGGAC 2416
Qy 471 LysGlyLeuGluThrAlaValAsnAspLeuIleSerProSerGlyAlaProLeu 490
Db 2417 CGTGGCGTGGAGAACACGTCGCGCTGACCTGTCTACCGCATGGTGGCGCGCATCTG 2476
Qy 491 IleSerAlaLeuGlyAspLeuGlyPheThrHisAlaAspValThrAlaValProSer 510
Db 2477 CTCAGCGCTCGCGCATATCGACGGCTTCGCGCATGACGACCTGACCGCGTG---CAG 2533
Qy 511 ThrIlePheThrSerProValPheThrThrGlyThrSerValAspTyrAlaGluLeuAsn 530
Db 2534 TTGCAGTACGCGCGCGCGCGCTGACCAATGGCAAGCATCATGTCGCGCGCGCGCGG 2593
Qy 531 ProSerIleIleValArgAlaGlySerPheAspProSerSerGlnProAsnAspArgHis 550
```

Db 2594 CCGCAGTGGTGGTGGCGCAGCGGTACCGTGGCGCGCCGCCCAACAAACAAATCCGCGC 2653
Qy 550 sValalaphSerThrAspGlyGlyLysAsnTrpPheGln---GlySerGluProGlyGlu 569
Db 2654 C---TGTATTACGGCTGTGGCGGCAAGCAATGCGTTCGTTGCGAGTGGCGCCAGC 2710
Qy 569 yValThrThrGlyThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaPr 589
Db 2711 CGGGCAGGGCCCGCAGCATTCGATTTGGCGATGCGCGCAGGTTGGTGGCGCAC 2770
Qy 589 oGlyAspProGlyGlnProValValTrpAlaValGlyPheGlyAsnSerTrpAlaAla 609
Db 2771 GCGCGCGCGCGCAAT-----TGGCGCAGCTCCGATTCGGCGCGAGTGGCAGCGTGT 2824
Qy 609 rGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThrPheTy 629
Db 2825 GAACGCCCTGCCGAACCCGCGTGGTGGTACCGATCGGTGGAGCAGCGCGCTGGTA 2884
Qy 629 rAla-----LeuSerAsnGlyThrPheTyArgSerThrAspGlyGlyValThrPheGlu 647
Db 2885 TCGCGTGGATGTCCGACGCGCGAGCTGTACGAGAGCACCAGTCCGCGCGCAGTTCCG 2944
Qy 647 nProValAlaAlaGlyLeu----- 653
Db 2945 T-----GCGACCGTGTACAGTGGCGCAGCGCGCGCGCGATGAGCGCACCCGCGCGCA 2998
Qy 654 -----ProSerSerGlyAlaValGlyValMetPheHisAlaValProGlyLysGluGlu 671
Db 2999 GCTGCGTCCGAGCCGCGCGCGAGCGGTGTATCTGCGCAGCCCGCGCAG----- 3053
Qy 671 yAspLeuTrpLeuAlaAlaSerSerGlyLeuTyHisSerThrAsnGlyGlySerTr 691
Db 3054 -----GGTGTGATCGCTGGCGAGCAGGT-----GCGCT 3082
Qy 691 pSerAlaIleThrGlyValSerSerAlaValAsnValGlyPheGlyLysSerAlaProGlu 711
Db 3083 CGAGTGTCTCCACAGCGCGGAGCAGCGTCTCGTGGCATCGCAAGCGCGTGGCGG 3142
Qy 711 ySerSerTyProAlaValPheValGlyThrIleGlyValThrGlyAlaTyAr 731
Db 3143 CGTGGCGCGCGCGCTGTATCTGGCGCGCGGTGCGAGCGGTGATGGCGTTCG 3202
Qy 731 gSerAspAspCysGlyThrTrpValLeuIleAsnAspAspGlnHisGlnTyGlyAs 751
Db 3203 CTCGAGAGTGGCGCGCGAGTGGCAGCGCATCACGATCAGCGCACCGCTTCGCGG 3262
Qy 751 nTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgValTyIleGlyThrAs 771
Db 3263 C---CCGTACAGCGTACCGGTGATCCGCGCATTCGCGCGGTGTGTACTTCGCGCACCG 3319
Qy 771 nGlyArgGlyIleValTyGlyAspIleGlyAlaProSerGlySerProSerProse 791
Db 3320 CGGCGCGCGCATTTTCACGCGCATCCGAGATGACGATGATGATCCGTCGCGA----- 3374
Qy 791 rValSerProSerAlaSerProSerLeuSerProSerProSerProSerProSerPro 810
Db 3375 ---CGTCTGCTCTTACCGTGTACGCGCGCGCGCGCTGGCGCGCTGGCGCGCGCTG 3430
Qy 811 -----SerProSerProSerProSerProSerProSerProSerProSerPro 827
Db 3431 ACACGCGCGACCTGGAGATCACCGCATGCCCGT-CCACATTTCCGCGCGCTGCTTG 3489
Qy 828 -----ProSerProSerProSerProSerProSerProSerProSerPro 836
Db 3490 CTGGTGTGTGTGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCTGCTGTT 3549
Qy 837 ----- 850
Db 3550 GCCGATGGCGGTGCTGCAGCGCGATGAGCAATGCGGTATGGCGTGGCGCGCGCGCC 3609
Qy 850 oSerProSerSer-----SerProSerSerProSerProSerProSerPro 864
Db 3610 GCGCGCGCAATCACGTTGAGTTTCGATGCGCAAGCAGCGCGCGCGCGCGCGCGCG 3661

RESULT 10

US-10-395-241-11
; Sequence 11, Application US/10395241
; Publication No. US20040038367A1
; GENERAL INFORMATION:
; APPLICANT: YAOI, Katsuro
; APPLICANT: MITSUISHI, Yasushi
; TITLE OF INVENTION: NOVEL XYLOGUCAN OLIGOSACCHARIDE-DEGRADING ENZYME, POLYNUCLEOTIDE
; FILE REFERENCE: 073756
; CURRENT APPLICATION NUMBER: US/10/395,241
; CURRENT FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: JP 2002-83433
; FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 2646
; TYPE: DNA
; ORGANISM: Geotrichum sp. M128
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (120)..(2558)
US-10-395-241-11

Alignment Scores:
Pred. No.: 2,646-69 Length: 2646
Score: 1125.50 Matches: 287
Percent Similarity: 49.8% Conservative: 131
Best Local Similarity: 34.2% Mismatches: 294
Query Match: 21.9% Indels: 127
DB: 7 Gaps: 32

US-09-917-376-1 (1-957) x US-10-395-241-11 (1-2646)

Qy 17 ValSerLeuLeuAlaAlaThrAlaSerPheAlaValAlaAlaAlaLeuGlyValLeuPro 36
Db 114 TTAGATATGGTGGCAGTCACCTCCCTCGGAAGCGCTTACTGCCCTTCGATTCTG--- 170
Qy 37 IleAlaIleThrAlaSerProAlaHisAlaAlaThrGlnProTyThrTrpSerAsn 56
Db 171 -----GCGTCGCTCGCGGTGCGCAGGAGCAGCTACGAGTTCAAGAAT 212
Qy 57 ValAlaIleGlyGlyGlyPheValAspGlyIleValPheAsnGluGlyAlaProGly 76
Db 213 GTCGCGATCGCGCGCGGTACATTACCGGATTTGTCGCGACCCCAAGACCAAGGAC 272
Qy 77 IleLeuTyValArgThrAspIleGlyMetTyArgTrpAspAlaAlaAsnGlyArg 96
Db 273 CTGCTGACGCGCGCAGCAGCATTTGGCGCGGTACCGTGGGACGCGCAGCTCCAG 332
Qy 97 TrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyTrpAsnGlyValValSer 116
Db 333 TGGATCCCGCTCAACGACTTTATCGAGGCGCAGGACATCAACATTATGGCAGCGAGTCG 392
Qy 117 IleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaValGlyMetTyThr--- 135
Db 393 ATCGCGCTGGAGCCCAACACCCCGAGGCTGTATCTCGCGCAGCGCGCTATGTCGC 452
Qy 136 AsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAspGlnGlyAlaThrTrp 155
Db 453 GACGAGTGG-----GCGCGTTCATGTGTCGAGACCGCGCGCGCTGTT 500
Qy 156 GlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGlu 175
Db 501 ACAATCTACGAGTCGCGCTTCCGATGGCGCCACGACGATGGAGCGCAACAATGGCGAG 560
Qy 176 ArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyThrPheGlyAlaProSerGlyLys 195
Db 561 GCGCTCGCTGTCAACCCCGTTCAACTCGAACGAGGTCTGGATGGGTACCGGTACA---GAG 617
Qy 196 GlyLeuTyArgSerThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPheProAsp 215

Db	618	GGTATCTGGAGAGTTTCGGACCGCGCAACACCTGGACAAACGTCACGTCATCCATCCCGGAC	677
Qy	216	ValGlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGlnSerAspIleGlnGly	235
Db	678	CGGTC-----ACNACGGTATCGGATACAGCTG-	707
Qy	236	ValValTrpValAlaPheAspLysSerSerSerLeuGlyGlnAlaSerLysThrIle	255
Db	708	-----GTCATTTTCGACCC-	740
Qy	256	PheValGlyValAlaAspProAsnAsnProValPheTrpSerArgAspGlyGlyAlaThr	275
Db	741	TACCGGAGCGAGCTGCCCGCGAGGGC--ATGTACGTCACGACACGCGCGGTGTCTCG	797
Qy	276	TrpGlnAlaValProGlyAlaProThrGlyPheIle-----	287
Db	798	TGGGAGCAGTGGGGCGAGCGTCGACGTGGCTCAACAGGACACGGGGCGGTTCCTCG	857
Qy	288	-----ProHisLysGlyValPheAspProValAsn	297
Db	858	GACAAGAAGCCCGCTGATCGCGCGCAGCCCATGAAGTCGTCTCAACCCCC-	911
Qy	298	HisValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyrAspGlySerSerGlyAsp	317
Db	912	AACCTCTCTACGTGACTTACGCGCGACTACCTCGTCCATGGGCGGTGTCAGTTCGGCGAA	971
Qy	318	ValTrpLysPheSerValTrpSerGlyThrTrpThrArgile-----	331
Db	972	GTCTGGCGCCAGAACCGCACCTCGGGCGCTGGAGCAGCATTACTCCCGCTCGGCAAC	1031
Qy	332	---SerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyrSerGlyLeuThr	350
Db	1032	TCGTGCGCTGCCCGTACAAACACAGAGCTTCCTCGGGCGGATTTTCGGGTCTCAGC	1091
Qy	351	IleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTrpTrpAsp	370
Db	1092	GTCGACGCGACCAACCCCAACCGTCTCGTGTCTATCACC--CTCAGCCGCGACCCCGGA	1148
Qy	371	ThrIle-----IlePheArgSerThrAspGlyGlyAlaThrTrpThrArgIleTrp	387
Db	1149	CCGCGCTCGACAGCATCTACCTCTCAACGATGCGCGCGACCTGGAGGACGTCACC	1208
Qy	388	AspTrpThrSer-----TyrProAsnArgSerLeuArg	398
Db	1209	CAGCTCTGTCGCCGCTCCACCTCGAAGGTAACTGGGGCCACCGACTACGGCGCGCG	1268
Qy	399	TyrValLeuAspIleSerAlaGluProTrpLeuThrPheGlyValGlnPro-----	415
Db	1269	TAC--AAGACGCGCACCGCTGTTCGTGCTCGACTTCAACACGCTCCCACTGGGG	1325
Qy	416	-----AsnProValProSerPro-----LysLeuGlyTrpMetAspGluAla	430
Db	1326	GGATAGTCGCGCGCAGCGGTACGCCCGGCTCACCAGTTTGGCTGTGGATGAGCGCT	1385
Qy	431	MetAlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThrGlyAlaThrLeuTyr	450
Db	1386	GTGCTTATCGATCGTTCAACCCCGACACCTGATGTACGCGCGGGGCGACCATCTGG	1445
Qy	451	AlaThrAsnAspLeuThrLys-----TrpAspSerGlyGlyGlnIleHisIle	466
Db	1446	GCACCGCACGCTCTCCCGTGTGAGAGGACTGG-----	1481
Qy	467	AlaPro-----MetValLysGlyLeuGluThrAlaValAsnAspLeuIle	482
Db	1482	GCSCCGAGCTGGTACCTCCAGATCGACGGTATCGAGGAGATCGATCTGTCTCGCTCCCG	1541
Qy	483	SerProProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHis	502
Db	1542	TCGCCCCAAGAGCGCGCGGCTCTGTCTGGGATCGGTGACATTAGCGCATGAGCAC	1601
Qy	503	AlaAspValThrAlaValProSerThrIlePheThrSerProValPheThrGlyThr	522

[illegible]

```

RESULT 11
US-10-395-241-17
; Sequence 17, Application US/10395241
; Publication No. US20040038367A1
; GENERAL INFORMATION:
; APPLICANT: YAOI, Katsuro
; APPLICANT: MITSUISHI, Yaeushi
; TITLE OF INVENTION: NOVEL XYLOGUCAN OLIG
; TITLE OF INVENTION: ENCODING THE SAME, A
; FILE REFERENCE: Q73756
; CURRENT APPLICATION NUMBER: US/10/395, 241
; CURRENT FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: JP 2002-83433
; PRIOR FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 19

```

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 17

; LENGTH: 2481

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Xyloglucan Oligosaccharide-Degrading Enzyme with Histidine Tag

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(2481)

US-10-395-241-17

Alignment Scores:

Pred. No.:	3,14e-69	Length:	2481
Score:	1124.00	Matches:	298
Percent Similarity:	48.9%	Conservative:	127
Best Local Similarity:	34.3%	Mismatches:	106
Query Match:	21.9%	Indels:	138
DB:	7	Gaps:	33

US-09-917-376-1 (1-957) x US-10-395-241-17 (1-2481)

QY	52	TyrThrTrpSerAsnValAlaIleGlyGlyPheValAspGlyValPheAsn	71
DB	13	TACGAGTTCAGAAAGTGTCCGATCGCGCGCGGGTACATTACCGGATTTCCGCGCAC	72
QY	72	GluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyMetTyrArgTyrAsp	91
DB	73	CCAAAGACCAAGGACCTGTGTACGCGCGCACGACATTTGGCGGCGGTACCGTGGAC	132
QY	92	AlaAlaAsnGlyArgTyrIleProLeuLeuAspTrpValGlyTyrAsnAsnTrpGlyTyr	111
DB	133	GCAGGACGTCCTCAAGTGTATCCCGTCAACGACATTTATCGAGGCGCAGACATGAACATT	192
QY	112	AsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaVal	131
DB	193	ATGGGACACGAGTGTGATCGCGCTGGACCCCAACACCCGACGAGCTGTACCTCGCGCAG	252
QY	132	GlyMetTyrThr---AsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAsp	150
DB	253	GGGCGCTATGTGGCGAGGAGTGG-----GCGGCGTTCATTGTCTCGAAGAC	300
QY	151	GlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGly	170
DB	301	CGCGGCGAGTGTATTACATCTACGAGTCCGCGTTCGATGGCGGCCAACGACATGGA	360
QY	171	ArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGly	190
DB	361	CGCAACAATGGCGAGCGCTCGCTGTCAACCCGTCAACTCGAACGAGGTCTGGATGGGT	420
QY	191	AlaProSerGlyLysGlyLeuTyrArgSerThrAspSerGlyAlaThrTrpSerGlnMet	210
DB	421	ACGCGTACA---GAGGATATCTGAAGAGTTCGACCGCGCCCAAGACCTGGACAAACGTC	477
QY	211	ThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThrGln	230
DB	478	ACGTCCATCCCGAGCGGCTC-----ACCAACGGTATCGATACACG	519
QY	231	SerAspIleGlnGlyValValTrpValAlaPheAspLysSerSerSerSerLeuGlyGln	250
DB	520	TCG-----GTCAATTTTCGACCCC-----GAA	540
QY	251	AlaSerLysThrIlePheValGlyValAlaAspProAsnAsnProValPheTrpSerArg	270
DB	541	CGTAATGGCACCATCTACGAGGCGGCGACTCCCGCGCAGGGC---ATGTACGTCAACGAC	597
QY	271	AspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPheIle-----	287
DB	598	GACGGCGGTGTCTCGTGGAGACGATGGCGGCGCAGCGCTCCAGCTGGCTCAACAGGAC	657
QY	288	-----ProHisLysGlyVal	292
DB	658	ACGGGCGGTTCCCGACAGAACGCCCGCTCGATCGCGCGCGGCGGCGGCGGCGGCGGCGG	717

QY	293	PheAspProValAsnHisValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyrAsp	312
DB	718	CTCACCCCC-----AACTTCTCTACGTGACTTACGCGGACTTACCTGCTGCTCCATGGGC	771
QY	313	GlySerSerGlyAspValTrpLysPheSerValThrSerGlyThrTrpThrArgIle---	331
DB	772	GTACGTTCCGCAAAAGTCTGGCGCGCAACCGGCTCGGCGGCGCTGGGACGACATTACT	831
QY	332	-----SerProValProSerThrAspThrAlaAsnAspTyrPheGly	345
DB	832	CCCCCGCTCGGCAACTCGTCCGCTCCCGGTACACACACGACAGCTTCCCTGCGGGCGGA	891
QY	346	TyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIle	365
DB	892	TTTTCGGGTCTCAGCGTCGACGCGACCAACCGTCTCGTCTCATCACC---CTC	948
QY	366	SerTrpTrpProAspThrIle-----IlePheArgSerThrAspGlyGlyAlaThr	382
DB	949	GACCGCGACCCCGGACCCCGCTCGACAGCATCTACCTCTCAACCGATGCGCGCGGACC	1008
QY	383	TrpThrArgIleTrpAspTrpThrSer-----TyrPro	393
DB	1099	TGAAAGGAGCTCACCAGCTCTCGTCCCGTCCAACTCGAAGTAACTGGGGCCACCCG	1068
QY	394	AsnArgSerLeuArgTyrValLeuAspIleSerAlaGluProTrpLeuThrPheGlyVal	413
DB	1069	ACTAAGCGCGCGGTAC---AAGGACGCGACGCGCTGTTCCGTGGCTCGACTTCAACAAC	1125
QY	414	GlnPro-----AsnProValProSerPro-----LysLeuGly	425
DB	1126	GGTCCCGAGTGGGGGATACGGTGGCGCGACGGTACGCCCGGCTCACCAAGTTGGC	1185
QY	426	TrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThr	445
DB	1186	TGGTGGATGAGCGCTGTGCTTATCGATCCGTTCAACCCGACGACCTGATGACGCGACG	1245
QY	446	GlyAlaThrLeuTyrAlaThrAsnAspLeuThrLys-----TrpAspSerGly	461
DB	1246	GGGGCGACCATCTGGGCGACCGACGCTCTCCCGTGTGAGAGGACTGG-----	1296
QY	462	GlyGlnIleHisIleAlaPro-----MetValLysGlyLeuGluThrAla	477
DB	1297	-----GCGCGAGCTGGTACCTCCAGATCGACGGTATCGAGGAGATGCG	1341
QY	478	ValAsnAspLeuIleSerProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeu	497
DB	1342	ATCTTGTCTCGCTCGCTCGCCCAAGAGCGCGCGCGCTCTCTGTCGGGCATCGGTGACATT	1401
QY	498	GlyGlyPheThrHisAlaAspValThrAlaValProSerThrIlePheThrSerProVal	517
DB	1402	ACGCGCATGAAGACGACGACGACCTCCACCAAG-----CCCCAAGATGTTGGTGGCGCCCG	1458
QY	518	PheThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAla	537
DB	1459	TCTTCCACCTCGACGATCGACGCTGGGCGCAACTTCCCAACGTTGCTGCTCGCGCC	1518
QY	538	GlySerPheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGly	557
DB	1519	GGATCTCGGACACGAGTACGACGCGCTCGCGCGCGGTGCGTACGCGACTGACGCGC	1578
QY	558	GlyLysAsnTrp-----PheGlnGlySerGluProGlyGlyValThr-----	572
DB	1579	GGAGCGGTGGACCATCTCTCCCTACCTGCCCTCTGGCATGAACCGGACCACTACCAG	1638
QY	573	GlyGlyThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaPro-----Gly	590
DB	1639	GACGACGATTTGAGTGCAGCGGCGGCGACCCAGATCGTGTGGTTCGACCAAGCTTGAC	1698
QY	591	AspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGln	610
DB	1699	GAGCAGGCGCTCGGACCGTGTGTTACTCGCACGACTATGGCAAGACGTGGTCT-----	1749


```
Qy 611 GlyValProAla-----AsnAlaGlnIleArgSerAspArgValAsn 624
Db 1750 ---GTTCCCGCTGGCGACTGAAGCCAGACTGCTCAATGTCTCGCAAGTCCAG 1806
Qy 625 ProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyVal 644
Db 1807 GATGGCAGTTCTACGTACCGATGGCGCAAGTTCTTCGTCTCGACCGACGGCGGAG 1866
Qy 645 ThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHis 664
Db 1867 TCGTATCCGCAAGAGCGCGCGACTTGTCACT-----GGCACATCGCTCATCGCTGCC 1920
Qy 665 AlaValProGlyLysGluGlyAspLeuTyrPhe---AlaAlaSerSerGlyLeuTyrHis 683
Db 1921 GTGAACCCCTGGTGGCGCGCGAGCTGTGGTGTCTGTCGCCAGGCGGTCTCTCCAC 1980
Qy 684 SerThrAsnGlyGlySerSerTyrPheAlaIle---ThrGlyValSerSerAlaValAsn 702
Db 1981 TCGACCGACTTTGGCGCTGTTACAGAGGTAGTACCGCAACGCGACCTCGTGAGC 2040
Qy 703 ValGlyPheGlyLysSer-----AlaProGlySerSerTyrProAlaValPhe 718
Db 2041 GTCGGCGCCCGCAAGTCCAAGTCGAGCGCAAGAGGCTAGCGCGCTCGCGGTCTTC 2100
Qy 719 ValValGlyThr-----IleGlyValThrGlyAlaTyrArgSerAspCysGly 736
Db 2101 ATTCGGGCGACCGCAAGCCCTGGAGCGACATCGGCTGTACCGCTCGACGACACCGC 2160
Qy 737 ThrThrTyrValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnTyrGlyGlnAlaIle 756
Db 2161 AGCACTGGACCGCGCTCAATGACGAGGACCAACTACTCGGGC---CCACCATGATC 2217
Qy 757 ThrGlyAspHisAlaAsnLeuArgValTyrIleGlyThrAsnGlyArgGlyVal 776
Db 2218 GAGSCCGACCCCAAGTCTACGGCGCGTGTATCTAGCAGCAACGCGCGGTATCGTG 2277
Qy 777 TyrGlyAspIleGlyAlaProSerGlySerProSerProSerProSerProSerAla 796
Db 2278 TACCGCGACCTTACCAACAAGAGAGC---AACGAGGAGAGTTCGACCGCAAGTGGCGC 2334
Qy 797 Ser-----ProSerLeuSer 801
Db 2335 AACGGCCAGAGGCGACGACTGTATGTGAAAAGAGATCTGGGTACCTGTGGCCAC 2394
Qy 802 ProSerProSerProSerProSerProSerProSerProSerProSerProSerProSe 821
Db 2395 CGGTTCCATGGCTGATATCGATCCGATTCGAGATTCGAGTCCGTCGACAGCTTGGCGCGCA 2454
Qy 821 rSerSer-ProSerProSerPro 828
Db 2455 CTCGAGCACCACCACCAACCA 2477

RESULT 12
US-10-395-241-13
; Sequence 13, Application US/10395241
; Publication No. US20040038367A1
; GENERAL INFORMATION:
; APPLICANT: YAOI, Katsuro
; APPLICANT: MITSUISHI, Yasushi
; TITLE OF INVENTION: NOVEL XYLOGLUCAN OLIGOSACCHARIDE-DEGRADING ENZYME, POLYNUCLEOTIDE
; FILE OF INVENTION: ENCODING THE SAME, AND METHOD OF PREPARING THE ENZYME
; FILE REFERENCE: Q73756
; CURRENT APPLICATION NUMBER: US/10/395,241
; CURRENT FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: JP 2002-83433
; PRIOR FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 2367
; TYPE: DNA
; ORGANISM: Geotrichum sp. M128
; FEATURE:
```

```
; NAME/KEY: mat_peptide
; LOCATION: (1)..(2367)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2367)
US-10-395-241-13
```

Alignment Scores:

```
Pred. No.: 9,28e-69 Length: 2367
Score: 1117.00 Matches: 281
Percent Similarity: 50.4% Conservative: 124
Best Local Similarity: 35.0% Mismatches: 279
Query Match: 21.8% Indels: 120
DB: 7 Gaps: 31
```

US-09-917-376-1 (1-957) x US-10-395-241-13 (1-2367)

```
Qy 52 TyrThrTyrSerAsnValAlaIleGlyGlyGlyPheValAspGlyIleValPheAsn 71
Db 10 TACGAGTTCAAGAATGTCCGATCGCGCGCGGGTACATTACCGGGATTGTCCGCGAC 69
Qy 72 GluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTyrAsp 91
Db 70 CCAAGACCAAGGACCTGCTGTACCGCGCGACGACATTGGCGCGGTACCGCTGGAC 129
Qy 92 AlaAlaAsnGlyArgTyrPheProLeuLeuAspTyrValGlyTyrAsnAsnTyrGlyTyr 111
Db 130 GCAGSCACGTCCAAGTGGATCCCGCTCAACGACTTTATCGAGCGCGCAGACATGAACATT 189
Qy 112 AsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTyrPalaIaVal 131
Db 190 ATGGCACCAGTCTCGATCGCGTGGACCCCAACAACCCCGCAGAGGTGTACCTCGCGCAG 249
Qy 132 GlyMetTyrThr---AsnSerTyrPaspProAsnAspGlyAlaIleLeuArgSerSerAsp 150
Db 250 GGGCGCTATGTCCGCGACGAGTGG-----GCGCGCTTCTATGTGTCGGAAGAC 297
Qy 151 GlnGlyAlaThrTyrPglInIleThrProLeuProPheLeuGlyGlyAsnMetProGly 170
Db 298 CGCGCCAGTGGTTTACATCTACGAGTCCGCTCCGATGGCGCCACGACATGGGA 357
Qy 171 ArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGly 190
Db 358 CGCAACAATGGCGAGCGCTCGCTGTCAACCGTTCACTCGAACGAGGTCTGGATGGGT 417
Qy 191 AlaProSerGlyLysGlyLeuTyrPaspSerThrAspSerGlyAlaThrTyrSerGlnMet 210
Db 418 ACGGTACACA---GAGGGTATCTGGAAGAGTTCGGACCGCGCCCAAGACCTGGACAACGTC 474
Qy 211 ThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGln 230
Db 475 ACGTCCATCCGCGACGCTTC-----ACCAACGGTATCGGATACAG 516
Qy 231 SerAspIleGlnGlyValTyrPalaPheAspLysSerSerSerSerLeuGlyGln 250
Db 517 TCG-----GTCATTTTCGACCCC-----GAA 537
Qy 251 AlaSerLysThrIlePheValGlyValAlaAspProAsnAsnProValPheTyrSerArg 270
Db 538 CGTAATGCGACCATCTACGAGCGCGCTGTCGCGCAGGGC---ATGTACGTTCACGCGAC 594
Qy 271 AspGlyGlyAlaThrTyrPglAlaValProGlyAlaProThrGlyPheIle----- 287
Db 595 GACGCGGTGTCTCTGTGGGAGCCAGTGGCGGCGCAGCTCCGCTGGCTCAACAGGACC 654
Qy 288 -----ProHisLysGlyVal 292
Db 655 ACGGGCGCTTCCCGCAAGAGCCGCGTGTATCGCGCGCGCGACCCATGAAGTCGCT 714
Qy 293 PheAspProValAsnHisValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyrAsp 312
Db 715 CTCACCCCC-----AACTCTCTACGTGACTTACGCGCGACTACCTCGTCCATGGGGC 768
```


QY 313 GlySerSerGlyAspValTrpLysPheSerValThrSerGlyThrTrpThrArgile--- 331
Db 769 GTCACGTTCCGGAAGTCTGGCGCAGAACCGACCTCGGGCGCTCGGAGACACATACT 828
QY 332 -----SerProValProSerThrAspThrAlaAsnAspTyrPheGly 345
Db 829 CCCCGGTCGCAACTCGTCGCTGCCCGTACAAACACAGAGCTTCCTCGGGCGGA 888
QY 346 TyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIle 365
Db 889 TTTTCGGCTCTCAGCGTCGACGGCGACCAACCAACCGCTCGTCTCATCACC---CTC 945
QY 366 SerTrpTrpProAspThrIle-----IlePheArgSerThrAspGlyGlyAlaThr 382
Db 946 GACCGGACCGCGGACCGCCCTCGACAGCATCTACCTCTCAACCGATGCCGGCGGAC 1005
QY 383 TrpThrArgIleTrpAspThrSer-----TyrPro 393
Db 1006 TGAAGGAGCTACCCAGCTCTGTCCTCCGTCGAACCTCGAAGTAACCTGGGGCCACCG 1065
QY 394 AsnArgSerLeuArgTyrValLeuAspIleSerAlaGluProTrpLeuThrPheGlyVal 413
Db 1066 ACTAAGCGGCGGCTAC---AAGACGGCACGCTGTCGTGCTCGACTTCAACAAC 1122
QY 414 GlnPro-----AsnProValProSerPro-----LysLeuGly 425
Db 1123 GGTCCCAAGTGGGGGGATACGCTGCGCGCACGCTACGCGCGGCTCACCAAGTTTGGC 1182
QY 426 TrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThr 445
Db 1183 TGTGTGATGACGCTGTGCTTATCGATCCGTTCACCCCGACGACCTGTATGACGGCAG 1242
QY 446 GlyAlaThrLeuThrAlaThrAsnAspLeuThrLys-----TrpAspSerGly 461
Db 1243 GGGGGACCATCTGGGCGACGACACGCTCTCCGTCGAGAGGACTGG----- 1293
QY 462 GlyGlnIleHisIleAlaPro-----MetValLysGlyLeuGluThrAla 477
Db 1294 -----GGCGGAGCTGGTACCTCCAGATCGAGCGGTATCGAGGAGATCG 1338
QY 478 ValAsnAspLeuIleSerProProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeu 497
Db 1339 ATCTGTGCTCGCTCGCCCAAGAGCGCGCGCTCTCTGTGCGGATCGGTGACATT 1398
QY 498 GlyGlyPheThrHisAlaAspValThrAlaValProSerThrIlePheThrSerProVal 517
Db 1399 AGCGCATGAACGACGACGACCTCACCAAG-----CCCAAGAGATGTTGGTGGCGCC 1455
QY 518 PheThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAla 537
Db 1456 TTCTCAACCTCGACAGATCGAGCTCGGGCAACTTCCCAACGTTGTCTCGCGCC 1515
QY 538 GlySerPheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGly 557
Db 1516 GGATCCTCGGACACGAGTACGACGCGCTGCGCGCGGTGCGTACGCGACTGACGCG 1575
QY 558 GlyLysAsnTrp-----PheGlnGlySerGluProGlyGlyValThr----- 572
Db 1576 GGAGCGGTGGACATCTTCCCTACCTCCCTCGCATGAACGCGGACCTACCCAG 1635
QY 573 GlyGlyThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaPro-----Gly 590
Db 1636 GGCAGCAGATTGACGTGCGAGCGGCGGACGAGATCGTGTGTCGACCAAGCTTGCAC 1695
QY 591 AspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGln 610
Db 1696 GAGCAGGCTCGGACCGTGTACTCGCAACGACTATGCAAGACGTGTCT----- 1746
QY 611 GlyValProAla-----AsnAlaGlnIleArgSerAspArgValAsn 624
Db 1747 ---GTTCCCGTGGCGACCTGAAGGCCCGACAGCTGCCAATGTGCTCTCGGACAAAGTCCAG 1803
QY 625 ProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyVal 644

Db 1804 GATGGACAGTTTACCTTACCGATGGCGCAAGTTCTTCTGCTCGACCGGCGGAAG 1863
QY 645 ThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHis 664
Db 1864 TCGTATGCGCCCAAGGGCGCGGACTTGTCACT-----GGCACATCGCTCATGCTGCC 1917
QY 665 AlaValProGlyLysGluGlyAspLeuTrpLeu---AlaAlaSerSerGlyLeuTyrHis 683
Db 1918 GTGAACCCCTGGGTGGCGCGGACGCTCTGGTGGCTGTTCCTCCGAGGCGGTCTCTCCAC 1977
QY 684 SerThrAsnGlyGlySerSerTrpSerAlaIle---ThrGlyValSerSerAlaValAsn 702
Db 1978 TCGACCGACTTGGCGGCTCGTTCACAGGGTAGGTACCGCCCAACCGACCTCGTGAGC 2037
QY 703 ValGlyPheGlyLysSer-----AlaProGlySerSerTyrProAlaValPhe 718
Db 2038 GTCGGGCGCCCAAGTCCAAAGTCGGACGCAAGAGCTAGCGCGCTCCCGGGTCTTC 2097
QY 719 ValValGlyThr-----IleGlyGlyValThrGlyAlaTyrArgSerAspCysGly 736
Db 2098 ATCTGGGCGACGCAAGGCTTGAAGCGACATCGGCTGTACCGCTCCGACGACACGCG 2157
QY 737 ThrThrTrpValLeuIleAsnAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIle 756
Db 2158 AGCACCTGACGCGCTCAATCACCAGGAGCACACTACTCGGGC---CCCACCATGATC 2214
QY 757 ThrGlyAspHisAlaAsnLeuArgValTyrIleGlyThrAsnGlyArgGlyIleVal 776
Db 2215 GAGCGCGCCCAAGTCTACGGGCGGTATCTAGGACGCAACGCGCGGTATCGTG 2274
QY 777 TyrGlyAspIle 780
Db 2275 TACGCGGACCTT 2286
RESULT 13
US-10-420-191-1
; Sequence 1, Application US/10420191
; Publication No. US20040067569A1
; GENERAL INFORMATION:
; APPLICANT: No. US20040067569A1ozymes Biotech, Inc.
; APPLICANT: Rey, Michael W.
; APPLICANT: Zaretsky, Elizabeth J.
; APPLICANT: Haas, Jeffrey A.
; TITLE OF INVENTION: Polypeptides Having Xyloglucanase Activity And Nucleic Acids
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 10210.200-US
; CURRENT APPLICATION NUMBER: US/10/420,191
; CURRENT FILING DATE: 2003-04-18
; PRIOR APPLICATION NUMBER: US 60/373,987
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 1
; LENGTH: 5698
; TYPE: DNA
; ORGANISM: Tricoderma reesei
US-10-420-191-1
Alignment Scores:
Pred. No.: 6, 19e-56 Length: 5698
Score: 940.00 Matches: 374
Percent Similarity: 30.1% Conservatives: 136
Best Local Similarity: 22.1% Mismatches: 315
Query Match: 18.3% Indels: 871
DB: 7 Gaps: 59
US-09-917-376-1 (1-957) x US-10-420-191-1 (1-5698)
QY 25 SerPheAlaValAlaAlaLeuGlyValLeuProfileAlaIleThrAlaSerProAla 44
Db 128 TCTCGAGTCTTCCCTTCTCTGCGGGCGGTC-----ATCCCTCGC 169

QY 45 HisAlaAlaThrThrGln----- 50
Db 170 CATGCTGCCCTTCTTCAGAGAGCTCAGAACGGGAACAGAACCCCGGAGTAGGGACGG 229
QY 51 -----ProTyrThrTrpSerAsnValAlaile--GlyGlyGlyGlyPheValAspGly 67
Db 230 GTACGCGGAAATTCATGGAAGACGTCAAGCTCGGCGGCGGCGGCGGCTTCTCCCGCGC 289
QY 68 IleValPheAsn-----GluGlyAlaPro----- 75
Db 290 ATCATCTTCCAAAGTACCTTCTTGCAGTTCGAGCGCGCGCGCGCGAAGCAGGGCGCG 349
QY 76 -----GlyIleLeuTyrValArgThrAspIleGlyGlyMet 87
Db 350 TAGTAGAAGGTTCCCAAGACAAAGCGTAGCATATGCAGAACAGATATTGGCGGCGTG 409
QY 88 TyrArgTrp----- 90
Db 410 TACGGCTCAAGGGTCTGTTTCCGCATCGTATACGTGTTGTCTATTAACCCCGCGAC 469
QY 91 -----AspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpVal----- 104
Db 470 ATGGCGGAGTTCGCGGAGCTCATGACCGCGCTCACGGATGGATTGCTGATAATGCC 529
QY 105 GlyTrp-----AsnAsnTrpGly----- 110
Db 530 GGCTGGCACAAAGCGGCTGCTGAGTACCTGGCGGCGAGTCCCTAACGACTATTACGG 589
QY 111 -----TyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLeuVal 127
Db 590 CCGACCGTTCCTGGGCGATCGAGCTGTGGTTCGCTTGTATCCGAGGACGATCAAAAGGTG 649
QY 128 TrpAlaAlaVal----- 131
Db 650 TATGCGCGAGTACCGCGTAGTCGGACACGCGAAGTACGCTGCTGCTAGTTTCCAC 709
QY 132 -----GlyMetTyrThrAsnSerTrpAsp----- 139
Db 710 ATACGCGCTCACGCGCATGTATACGAACAGCTGGTC-TGTGATGTCCTCAGATCTAGACC 768
QY 139 ----- 139
Db 769 TATGATTGGACGCGGTACATATGCTTGTTCGACACAGACACTACAGAGAGTCTAGATCTGG 828
QY 140 -----ProAsnAspGlyAlaIleLeuArg 147
Db 829 ATACTAACCTGCCTGACATTGGCCATATATAGGATCCGAGTAAATGGAGCCATTCGCG 888
QY 148 SerSerAspGln----- 151
Db 889 TCGTCAGACGCGACTGTAAACCGGTATATATCCCTAGGCTCAITTAACCTCGGTAGTAAGCG 948
QY 152 -----GlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsn 167
Db 949 AGCAGTCTGGCGGCGGCAACGTGCTCTCCACCAACTTGGCCCTTCAAGTTCGGGGTAAC 1008
QY 168 MetProGlyArg----- 171
Db 1009 ATCCAGAGCGCGCGGTGCACAGGAAGTGGTTGAACGGGAAGTTTCAGCCCCCATATG 1068
QY 172 -----GlyMetGlyGlyArgLeuAlaValAspProAsnAsnAspAsnIleLeu 187
Db 1069 TACGGTCTCGGGAGCGCGAGAGCGTCTGGTGTGATCCGCGCAACTCCAACATCATC 1128
QY 188 TyrPheGlyAlaPro----- 192
Db 1129 TACTTGGTCTCTCGGCCCTCTCGCAGACCGACAGCTAGGCGCGTGGAGTTGTAGTAG 1188
QY 193 -----SerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrp 207
Db 1189 ATGAACACCGACGCTCAGGAACCGCCCTCTGGAAGTCTACGACGCGCGCGTACCTTT 1248
QY 208 SerGlnMet----- 210

Db 1249 TCCAAGGTCTCGGCGAGTCTTTTCCCGGAGACCTTTCAGATGCTTCCGCGCACTGGAAA 1308
QY 211 -----ThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThr 227
Db 1309 AGGTTCCAGAGCTGTTTCCCGCAACTGGGAGCTACATCCAGACCCGAGTATTCACAC 1368
QY 228 GlyTyrGln----- 230
Db 1369 GGCTACAACAGCAGCAAGTGCCTGGTTCGATGTAGGGTCTGGGCTCACTAAGGTTG 1428
QY 231 -----SerAspIleGlnGlyValValTrpValAlaPheAspLysSerSerSer 247
Db 1429 CCGATGTTGTCCGACAAAGGACTCATGTGGTTAGCTTTCGACTCAACAGCAGCAGC 1488
QY 248 LeuGlyGlnAla----- 251
Db 1489 ACCGGGGAGCCCTGTTCTGTTCTGAGTACACCCNAATGCAAGCTGAGTGGTGGTCGTCG 1548
QY 252 -----SerLysThrIlePheValGlyValAlaAspProAsnAsnProValPhe 267
Db 1549 TGGCCCCCTCGGACGCTCTCGTATCTTTGTTGGCAGCGCTGATAACATCACTGCTTCAGTC 1608
QY 268 Trp----- 268
Db 1609 TATGTGACACGTGCAGAGCATAGAAACACCGTCCGACTATTGTAGTGAAGAGTCAAG 1668
QY 269 -----SerArgAspGlyGlyAlaValTrpGlnAlaValProGlyAlaProThrGlyPhe 286
Db 1669 ATACACTGTGCAATCGCGCTCCACGTGGAGTGTCTGTACCGGGCAGCCAGGAAATAC 1728
QY 287 IleProHis----- 289
Db 1729 TTTCTCTACAGTTACGGCCGAGGTGCACCTCACGACATGCCCCGTCGTCCTTTATG 1788
QY 290 LysGlyValPhe-----AspProValAsnHisValLeuTyrIleAlaThrSerAsn 306
Db 1789 AAAGAGTGTTCGGGAAACTGCAGCCAGCAGAGAAGGCTTGTATCTGACCTATTCCGAT 1848
QY 307 ThrGlyGlyPro----- 310
Db 1849 GGCACAGGCGCGCGCTTTGACGTGCTGCTCTTCCGGAACATAGACTGGATAAGCTTA 1908
QY 311 -----TyrAspGlySerSerGlyAspValTrpLysPheSerValThrSerGly 326
Db 1909 CCGTGTCCCGCTATGATGGCACACTGGCTCAGTGTGGAGGTAGACATTCGACGGGGA 1968
QY 327 ThrTrpThrArgIleSerProValProSerThrAspThr----- 339
Db 1969 ACTTGGAAAGACATACTACCGTGTGAACCGAGTCACACCTCCATGCTGTAACGTCGCCCT 2028
QY 340 -----AlaAsnAspTyrPheGlyTyrSerGlyLeu 349
Db 2029 TGAACCTTTCTGATCACCCTGCTCTCTGGATCAGATCTATATCTTTGGCTTTGGCGCCT 2088
QY 350 ThrIleAsp----- 352
Db 2089 GGCCTCGATTGTAGTGGGACAGAGACCTTAGTATATGAACCGGAACCCGCGGAA 2148
QY 353 -----ArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTrpTrpPro 369
Db 2149 CCGGAGCTAAACCAAGCCAGGAACCTTGTGTGTCTTCTTCAACTCTTGGTGCCA 2208
QY 370 AspThr----- 371
Db 2209 GATGCTCAGTGGTGTTCGTCCTTGGGAACAACAACGAAGAACTTGAGAACCACCGGT 2268
QY 372 -----IleIlePheArgSerThrAspGlyAlaThrTrpThrArgIleTrpAspTrp 389
Db 2269 CTACGATCGCACTTTCGTCGACCGACTCTGGGCAACATCGAGCCCATCTGGCGGTGG 2328
QY 390 ThrSerTyrProAsnArgSer---LeuArgTyrValLeuAspIleSerAlaGluPro--- 407

Qy 741 -----LeuIleAsnAspGlnHisGlnTyrGly--- 750
Db 4527 GAGATCGGTCACTCTGTGCGCGGAGACCTCGGACATCCAGGCTCCAGGGTTC 4586
Qy 750 ----- 750
Db 4587 GGCTCCATCGACAGCACCAAGTTCGCGCGGAGCGGCTGTAGTCCCGAGGGTCCCGAAG 4646
Qy 751 -----AsnTrpGlnAlaIleThrGlyAspHisAlaAsnLeuArgVal 766
Db 4647 CCGAGGTAGTGTCTGTTCAGCGCGCTCGCGCAGCACCGCGGG-----CAAGTC 4700
Qy 767 TyrIleGlyThrAsnGlyArgGlyIleValTyrGlyAsp----- 779
Db 4701 TACGTGGGACCAACCGCGCGGGCTCTTTTACGCTCAGGCGTGTGCGCGCGCTTCAG 4760
Qy 780 -----IleGlyGlyAlaPro 784
Db 4761 ATGACCCGTGTTCGCGCGCGCGCAGAAATGCGAGTCCGACCGTTCGCGCGCGC--- 4817
Qy 785 SerGlySerProSerProSerValSerProSerAlaSerProSerLeuSerProSerPro 804
Db 4818 ACGGGGGGACTTCTCGTCGACCAACGACGAGCAGCAGTACTTGGCAGCGCGCGC 4877
Qy 805 SerPro-----SerSerSerProSerProSerProSerPro 816
Db 4878 TGCCCGCGCTGAAGAGCAGCTGCTCGTCTCGTGTCTGTCATCTCTTCGCGCAGCTCG 4937
Qy 817 SerSerSerProSerSerProSerProSerPro----- 828
Db 4938 AGCACCACGCTGAGGTGAGCGTGTATTCACGACCGCGGCTTGGAGAACGCGGTGAGC 4997
Qy 829 -----SerProSerProSer----- 833
Db 4998 TCGTGTGCGACTCCAGCTCGACATAGTGTGCGCGCGCGCGGAGCGGTGACTTCGTCG 5057
Qy 834 -----ProSerArgSerProSerProSerAlaSerProSerProSerSerProSerPr 852
Db 5058 AGGACCACTCGCGCGCGCGTCCACCGGGTTCAGGGGTTCGCGGCTGCCACTGAAGCAGC 5117
Qy 852 oSerSerSerProSerSerProSerProSerProSerProSerProSerProSer 865
Db 5118 TCCTGTGTCGAGCGCGCGCGCGGAGGTGCCCGCGCGCTCCCGCAGC 5157

RESULT 14

US-09-917-384-2
; Sequence 2, Application US/09917384
; Publication No. US20030096342A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOSE TOLERANT
; FILE REFERENCE: 40170.6US01
; CURRENT APPLICATION NUMBER: US/09/917,384
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 3687
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Segment of
; OTHER INFORMATION: GuxA
US-09-917-384-2

Alignment Scores:

Pred. No.: 6.43e-48 Length: 3687
Score: 823.00 Matches: 251

Percent Similarity: 47.1% Conservative: 46
Best Local Similarity: 39.8% Mismatches: 173
Query Match: 16.0% Indels: 164
DB: 3 Gaps: 24
US-09-917-376-1 (1-957) x US-09-917-384-2 (1-3687)
Qy 426 TrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThr 445
Db 307 TGGATGACCGCATCGCTGCGATCAACGGCGTCAACGGC-----GGACCC 351
Qy 446 GlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlnIleHis 465
Db 352 GGCTTGACGACATATCTGCGACCGCGCTCTCCAG-----CAGCAGGGA 396
Qy 466 IleAlaProMetValLysGlyLeuGluThrAlaValAsnAspLeuSerProPro 485
Db 397 ACACCCCTGAAGTC-----ATTGAGATTGTCTACTACGATCTG-----CGGGA 441
Qy 486 SerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspVal 505
Db 442 CGCGACTGCGCGCGCTCGCTCCAAACGGCGACTGCCCGCT---ACGGCAGCAGGTTTG 498
Qy 506 ThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSerValAsp 525
Db 499 CAGACCTAT---GAAACGCGAGTACATCGATCCGATTGCGAGTATCTCGACCAATCCGAAG 555
Qy 526 TyrAlaGluLeuAsnProSerIleIleValAlaGlySerPheAspProSerSerGln 545
Db 556 TACTCCAGCTGCGGATCGTGCAGCATCAIT-----GAGCCGAGCTCGCTG 600
Qy 546 ProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGlySer 565
Db 601 CCAACGC-GGTAC-----CAATAT 620
Qy 566 GlnProGlyGlyValThrThrGlyGlyThrValAla-----AlaSerAlaAspGly 582
Db 621 GAGCATTCACAGCGTGTCAACGGCGGTCCGCTATT-ACGAGCAGGCGATCGAGTACGCGC 679
Qy 583 SerArg-----PheValTrpAlaProGlyAspPro--- 592
Db 680 TCAGAAATTCACGCCCATTCGAAACGTTGATCATCTACATGAGCGCCGCCCTCCGCT 739
Qy 593 -----GlyGlnProValVal-----TyrAlaValGlyPheGlyAsnSerTrp 606
Db 740 GGCTTGGTGGCCCAATATCCAGCGGATACGTACAGAAAGTCCAGAGGTCTCTCAACG 799
Qy 607 AlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsn----- 624
Db 800 CGAGCATCGGGTCAACGGCATCGACGGCTTCGTCACCAACACACGGCGAATTACACGCCGT 859
Qy 625 -----ProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyr 637
Db 860 TGAAGGAGCGCTTCATGACCGCCACCCAGCAGGTGCGCGGACAGCGGTGGAGTCGGGA 919
Qy 638 ArgSerThrAspGlyGlyValThrPheGlnPro----- 648
Db 920 ATTCTACAGTGAATCTCTGACATC-GACGAAGCCGACTACGCGGTGACTGTGTACTCG 978
Qy 649 -----ValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHisAlaVal 666
Db 979 CGGCTCGTCCGCGTGGCTTCCAGAGCAGC-----ATCCGATGCTCATCGACACCTTA 1032
Qy 667 ProGlyLysGluGlyAspLeuTrpLeuAlaSerSerGlyLeuTyrHisSerThrAsn 686
Db 1033 -----CGCAAC 1038
Qy 687 GlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnVal----- 703
Db 1039 GGTGGGGTGTTCGGAACGAAACCAACAGCGCCGAGCAGCGGATGTCAACACCTTC 1098
Qy 704 -----Gly 704

1099 GTCAACAGTGAAGATTGACCTTCGGCAGCAGCCGCGCTGTGTGTGCAACAGACGAT 1158
 QY PheGlyysrAlaProGlySerTyrProAlaValPhe----- 718
 Db CGGGCTCGCCAGCCGCGCAGCAGCCGCGCAGCTTCGCGAGCGCAGCCTCGAC 1218
 QY ValValGlyThrIleGlyGlyValThrGlyAla----- 729
 Db CGGTATGTCTGGATCAAGCCCGCGGTGAGTTCGCGAGCGCAAGCGTGGAGGATCCG 1278
 QY TyrArgSerAspAsp---CysGlyThrThrTrpValLeuIleAsnAspAsp 745
 Db ACAACTGGCAAGAGTCGACCCCTGTCGCGACCGAGTACACGAG----- 1326
 QY GlnHisGlnTyrGlyAsnTrpGlyGlnAla-----ileThrGlyAspHis 760
 Db TCGTACGGGTACTGACCAAGCGGTACCGAACTCCCGATCGCGCGCAGTGG 1380
 QY AlaAsnLeuArgValTyrIleGlyThrAsnGlyArgGlyIleValTyrGlyAspIle 780
 Db TTCGCGCGCAGTGTGACCACTTGTTCGCGAAGCAGCAGCGCGTGCAGACGTGCGACC 1440
 QY GlyGly-----AlaProSerGlySerProSerProSerValSer 793
 Db AGCTCGAGCGCGCGCTCCGCGCGCGAGTCCGTCGCTTCGCGAGTCCGAGCCGAGT 1500
 QY ProSerAlaSerProSerLeuSerProSerProSerProSerProSerProSer 813
 Db CCGAGCGCGAGCAGTCCGATCGCGTCCGCTCCGAGCTCGAGCCGCTCCGCTCG 1560
 QY Pro-----SerProSerSerProSerProSerProSer 825
 Db CCGAGCGCGAGTCCGAGCGCGAGTCCGCGTCCGCTTCGCGAGTCCGAGCCGCT 1620
 QY ProSerProSerProSerProSerProSerProSerProSerProSerProSer 845
 Db CCGTCCGCGAGCGCGAGTCCGAGCGCGAGTCCGCGTCCGCTTCGAGTCCGAGC 1680
 QY ProSerSerProSerProSerProSerProSerProSerProSerProSer 865
 Db CCGTCTCGTCCGCGAGCGCGAGTCCGAGCGCGAGTCCGCTTCGCGTCCGCGAG 1740
 QY SerSerProValSerGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGly 885
 Db TCGTCCGCGGTGTGGTGGCTGAAGTGCAGTACAGAACATGATTCGCGCGCGGT 1800
 QY AsnAsnGlnLysProGlyLeuGlnValValAsnThrGlySerSerValAspLeu 905
 Db GATAACAGATCAACCGCGGTCTCCAGTTGTGTAATACCGCGTCTGTCGCTGGATTG 1860
 QY SerThrValThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyr 925
 Db TCGACGCTGAGCGGTGCGGTACTGTTTCCCGGATGTTGGTTCGACACTGGTGTAC 1920
 QY AsnCysAspTrpAlaAlaIleGlyCysGlyAsnIleArgAlaSerPheGlySerValAsn 945
 Db AACTGTACTGGCGCGGATGGGTGGTGGATATCCGCGCTCTGTCGCTCGGTGAC 1980
 QY ProAlaThrProThrAlaAspThrTyrLeuGln 956
 Db CCGCGAGCGCGCGCGGACACCTACTCTGAC 2013

RESULT 15

US-09-917-383-2
 ; Sequence 2, Application US/09917383
 ; Publication No. US20030104522A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DING, SHI-YOU
 ; APPLICANT: ADNEY, WILLIAM S.
 ; APPLICANT: VINZANT, TODD B.
 ; APPLICANT: DECKER, STEPHEN R.
 ; APPLICANT: HIMMEL, MICHAEL E.
 ; TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS

1099 GTCAACAGTGAAGATTGACCTTCGGCAGCAGCCGCGCTGTGTGTGCAACAGACGAT 1158
 QY PheGlyysrAlaProGlySerTyrProAlaValPhe----- 718
 Db CGGGCTCGCCAGCCGCGCAGCAGCCGCGCAGCTTCGCGAGCGCAGCCTCGAC 1218
 QY ValValGlyThrIleGlyGlyValThrGlyAla----- 729
 Db CGGTATGTCTGGATCAAGCCCGCGGTGAGTTCGCGAGCGCAAGCGTGGAGGATCCG 1278
 QY TyrArgSerAspAsp---CysGlyThrThrTrpValLeuIleAsnAspAsp 745
 Db ACAACTGGCAAGAGTCGACCCCTGTCGCGACCGAGTACACGAG----- 1326
 QY GlnHisGlnTyrGlyAsnTrpGlyGlnAla-----ileThrGlyAspHis 760
 Db TCGTACGGGTACTGACCAAGCGGTACCGAACTCCCGATCGCGCGCAGTGG 1380
 QY AlaAsnLeuArgValTyrIleGlyThrAsnGlyArgGlyIleValTyrGlyAspIle 780
 Db TTCGCGCGCAGTGTGACCACTTGTTCGCGAAGCAGCAGCGCGTGCAGACGTGCGACC 1440
 QY GlyGly-----AlaProSerGlySerProSerProSerValSer 793
 Db AGCTCGAGCGCGCGCTCCGCGCGCGAGTCCGTCGCTTCGCGAGTCCGAGCCGAGT 1500
 QY ProSerAlaSerProSerLeuSerProSerProSerProSerProSerProSer 813
 Db CCGAGCGCGAGCAGTCCGATCGCGTCCGCTCCGAGCTCGAGCCGCTCCGCTCG 1560
 QY Pro-----SerProSerSerProSerProSerProSer 825
 Db CCGAGCGCGAGTCCGAGCGCGAGTCCGCGTCCGCTTCGCGAGTCCGAGCCGCT 1620
 QY ProSerProSerProSerProSerProSerProSerProSerProSerProSer 845
 Db CCGTCCGCGAGCGCGAGTCCGAGCGCGAGTCCGCGTCCGCTTCGAGTCCGAGC 1680
 QY ProSerSerProSerProSerProSerProSerProSerProSerProSer 865
 Db CCGTCTCGTCCGCGAGCGCGAGTCCGAGCGCGAGTCCGCTTCGCGTCCGCGAG 1740
 QY SerSerProValSerGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGly 885
 Db TCGTCCGCGGTGTGGTGGCTGAAGTGCAGTACAGAACATGATTCGCGCGCGGT 1800
 QY AsnAsnGlnLysProGlyLeuGlnValValAsnThrGlySerSerValAspLeu 905
 Db GATAACAGATCAACCGCGGTCTCCAGTTGTGTAATACCGCGTCTGTCGCTGGATTG 1860
 QY SerThrValThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyr 925
 Db TCGACGCTGAGCGGTGCGGTACTGTTTCCCGGATGTTGGTTCGACACTGGTGTAC 1920
 QY AsnCysAspTrpAlaAlaIleGlyCysGlyAsnIleArgAlaSerPheGlySerValAsn 945
 Db AACTGTACTGGCGCGGATGGGTGGTGGATATCCGCGCTCTGTCGCTCGGTGAC 1980
 QY ProAlaThrProThrAlaAspThrTyrLeuGln 956
 Db CCGCGAGCGCGCGCGGACACCTACTCTGAC 2013

Alignment Scores:
 Pred. No.: 6,43e-48 Length: 3687
 Score: 823.00 Matches: 251
 Percent Similarity: 47.1% Conservatives: 46
 Best Local Similarity: 39.8% Mismatches: 173
 Query Match: 16.0% Indels: 164
 DB: 3 Gaps: 24

US-09-917-376-1 (1-957) x US-09-917-383-2 (1-3687)

QY 426 TrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThr 445
 Db 307 TGGATGGACCGCATCGCTCGCATCAACGGCGCTCAACGGC-----GGACCC 351
 QY 446 GlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlnIleHis 465
 Db 352 GCGTTGACGACATATCTGGACCGCGCTCTCCCCAG-----CAGCAGGGA 396
 QY 466 IleAlaProMetValLysGlyLeuGluThrAlaValAsnAspLeuIleSerProPro 485
 Db 397 ACCACCCCTGAAGTC-----ATTGAGATTGTCATCTACGATCTG-----CCGGGA 441
 QY 486 SerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspVal 505
 Db 442 CCGGACTCGCGCGCGCTCGCTCCAAACGGCGAAGTCCCGCT---ACGGCAGCAGGTTG 498
 QY 506 ThrAlaValProSerThrPheThrSerProValPheThrThrGlyThrSerValAsp 525
 Db 499 CAGACCTAT---GAAACCGCAGTACATCGATCGATTCGATTCGAGTATCTCGACAAATCCGAAG 555
 QY 526 TyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspProSerSerGln 545
 Db 556 TACTCCAGCCTCGCGATCGTGCATCAT---GAGCCGGACTCGCTG 600
 QY 546 ProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGlySer 565
 Db 601 CCAACGCG-GGTCAC-----CAATAT 620
 QY 566 GluProGlyGlyValThrThrGlyGlyThrValAla-----AlaSerAlaAspGly 582
 Db 621 GAGCATTCAGCGTGTGCAACGGCGTCCGCTATT-ACGAGCAGGATCATCGATACGCGC 679
 QY 583 SerArg-----PheValTrpAlaProGlyAspPro--- 592
 Db 680 TCACGAAATTCGACGCCATTTCGAAACGCTGTACATCTACATGACGCGCCCATCCGGCT 739
 QY 593 -----GlyGlnProValVal-----TyrAlaValGlyPheGlyAsnSerTrp 606
 Db 740 GCGTTGGGTGGCCCAATAATATCCAGCGGATACGAGGATCCAGAGTCTCTCAACG 799
 QY 607 AlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsn----- 624
 Db 800 CGAGCATCGGGTCAACGGCATCGCGCTTCGTCACCAACACCGCGGAATTACACGCCGT 859
 QY 625 -----ProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyr 637
 Db 860 TGAAGGAGCGGTTCATGACCGCCACCCAGCAGGTTCGCGCGGACGAGCCGCTGAGTTCGCGCA 919

; TITLE OF INVENTION: CELLULOYTICUS
 ; FILE REFERENCE: 40170.6US01
 ; CURRENT APPLICATION NUMBER: US/09/917,383
 ; CURRENT FILING DATE: 2001-07-28
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1.
 ; SEQ ID NO 2
 ; LENGTH: 3687
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Segment of
 ; OTHER INFORMATION: Guxa
 ; US-09-917-383-2

```
QY 638 ArgSerThrAspGlyGlyValThrPheGlnPro----- 648
Db 920 ATTCTACCAAGTGGAAATCCCTGACATC-GACGAAGCCGACTACGCGGTGACTTGTACTCG 978
QY 649 -----ValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHisAlaVal 666
Db 979 CGGCTCGTCCGCTGCTTCCAGAGC-----ATCGCATGCTCATCGACACCTTTA 1032
QY 667 ProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyHisSerThrAsn 686
Db 1033 -----CGCAAC 1038
QY 687 GlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnVal----- 703
Db 1039 GGTGGGGTGGTCCGAACGAACCAACAGGCCGAGCACCGCGACCGCATGTCAACACCTTC 1098
QY 704 -----Gly 704
Db 1099 GTCAACCAAGTCGAAGATTGACCTTCGGCAGCACCGCGGCTGTGGTGAACCAACGACGGT 1158
QY 705 PheGlyLysSerAlaProGlySerSerTy-ProAlaValPhe----- 718
Db 1159 GCGGGCTCGCCAGCGCGCGCAGCGAAGCCCGACGACTTCCCGAAGCGCGCCTCGAC 1218
QY 719 -----ValValGlyThrIleGlyValThrGlyAla----- 729
Db 1219 GCGTATGCTCTGGATCAAGCGCGCGGTGAGTCGGACGGCACAGCGCTGCGAGCATCCG 1278
QY 730 -----TyrArgSerAspAsp---CysGlyThrThrTrpValLeuIleAsnAspAsp 745
Db 1279 ACACTGGCAAGAGTCGACCCCATGTGCGACCCGACCGCTACACGACG----- 1326
QY 746 GlnHisGlnTyArgAsnTrpGlyGlnAla-----IleThrGlyAspHis 760
Db 1327 -----TCGTACGGGGTACTGACCAACGCGTTACCGAACTCCCGATCGCGGCCAGTGG 1380
QY 761 AlaAsnLeuArgValTyIleGlyThrAsnGlyArgGlyIleValTyGlyAspIle 780
Db 1381 TTCGGCGCGAGTTTGACCACTGTTCGCGAAGCAGCGCGCGGTGCGGAGCTCGAC 1440
QY 781 GlyGly-----AlaProSerGlySerProSerProSerValSer 793
Db 1441 AGCTCGAGCCCGCGCTCGCGCGCGAGTTCGCGGCTTCGCGAGTCCGAGCCCGAGT 1500
QY 794 ProSerAlaSerProSerLeuSerProSerProSerProSerProSerProSer 813
Db 1501 CCGAGCCCGAGCAGCTCGCCATCGCCGCTCGCGCTCCGAGCTCGAGCCCGCTCCGCTCG 1560
QY 814 Pro-----SerProSerSerProSerSerProSer 825
Db 1561 CCGAGCCCGAGTCCGAGCCCGAGTAGCTCGCGCTCGCGCTCCGAGCTCGAGCCCGCTCT 1620
QY 826 ProSerProSerProSerProSerProSerProSerProSerProSerAlaSerProSer 845
Db 1621 CCGTCGCGAGCCCGAGTCCGAGCCCGAGTAGCTCGCGCTCGCGCTCCGAGCTCGAGC 1680
QY 846 ProSerSerProSerProSerProSerProSerProSerProSerProSerProSer 865
Db 1681 CCGTCTCCGTCGCGCGAGCCCGAGTCCGAGCCCGAGTAGCTCGCGCTCGCGCTCCGAGC 1740
QY 866 SerSerProValSerGlyGlyValIysValGlnTyLysAsnAsnAspSerAlaProGly 885
Db 1741 TCGTCGCGGCTGTGCGGTGGGCTGAAGTGCAGTACAGAACAAATGATTTCGGCGCCGGT 1800
QY 886 AspAsnGlnIleLysProGlyLeuGlnValValAsnThrGlySerSerSerValAspLeu 905
Db 1801 GATAACCAAGATCAACCGGGTCTCCAGTTGTTGAATACCGGGTCTGTCGTTGATTG 1860
QY 906 SerThrValThrValArgTyTrpPheThrArgAspGlyGlySerSerThrLeuValTy 925
Db 1861 TCGACGGTACGGTGGCTGCTGCTTCCACCGGGATGGTGGTCTCGACACTGGTGTAC 1920
QY 926 AsnCysAspTrpAlaAlaIleGlyCysGlyAsnIleArgAlaSerPheGlySerValAsn 945
```

Search completed: March 2, 2006, 17:16:06
Job time : 15122 secs

Db 1921 AACTGTGACTGGGCGGCGATGGGGTGTGGGAATATCCCGCCTCGTTCGGCTCGGTGAAC 1980
QY 946 ProAlaThrProThrAlaAspThrTyLeuGln 956
Db 1981 CCGGCGACCGCGACCGCGGACACCTACCTGCGAG 2013

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocollaboration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 2, 2006, 12:24:59 ; Search time 1026.96 Seconds
(without alignments)
2043.020 Million cell updates/sec

Title: US-09-917-376-1

Perfect score: 5134

Sequence: 1 MDRENTLRLWRSRLVLL.....RASFGSVNPATPTADTYLQX 957

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7218512 seqs, 1096188070 residues

Total number of hits satisfying chosen parameters: 14437024

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framer_p2n.model -DEV=xlp
-O=/abss/ABSSWEB/spool/US09917376/runat_02032006_091502_8293/app_query.fasta.1
-DB=PublishedApplicationsNA.New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs03p
-USER=US09917376 @CGN_1_1067@runat_02032006_091502_8293 -NCPU=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7.
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.New.*

1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	282.5	5.5	9903	12	US-11-052-554A-517 Sequence 517, App
2	273	5.3	9474	12	US-11-052-554A-526 Sequence 526, App
3	273	5.3	11151	12	US-11-052-554A-525 Sequence 525, App
4	270.5	5.3	6615	12	US-11-052-554A-518 Sequence 518, App

5	264.5	5.2	6210	12	US-11-056-470-1	Sequence 1, Appl
6	264.5	5.2	7572	12	US-11-052-554A-527	Sequence 527, App
7	261.5	5.1	70513	8	US-10-995-561-13368	Sequence 13368, A
c 8	260	5.1	4146	12	US-11-052-554A-522	Sequence 522, App
c 9	258.5	5.0	88421	12	US-11-205-109-1	Sequence 1, Appli
c 10	250	4.9	3990	12	US-11-052-554A-520	Sequence 520, App
c 11	250	4.9	114801	12	US-11-121-086-22	Sequence 22, Appl
c 12	247.5	4.8	3240	12	US-11-052-554A-529	Sequence 529, App
c 13	247.5	4.8	7231	12	US-11-136-527-2622	Sequence 2622, Ap
c 14	247	4.8	6360	12	US-11-056-470-2	Sequence 2, Appli
c 15	243	4.7	2352	12	US-11-052-554A-541	Sequence 541, App
c 16	240.5	4.7	215126	7	US-10-330-773-339	Sequence 339, App
c 17	240	4.7	3000	12	US-11-136-527-2740	Sequence 2740, Ap
c 18	238.5	4.6	2845	9	US-11-072-512-945	Sequence 945, App
c 19	235.5	4.6	15876	12	US-11-052-554A-560	Sequence 560, App
c 20	235.5	4.6	200628	12	US-11-121-086-62	Sequence 62, Appl
c 21	234	4.6	3580	8	US-10-995-561-87	Sequence 87, Appl
c 22	233.5	4.5	173602	12	US-11-121-086-25	Sequence 25, Appl
c 23	232.5	4.5	2268	8	US-10-517-939-329	Sequence 329, App
c 24	232.5	4.5	63984	12	US-11-121-086-26	Sequence 26, Appl
c 25	230.5	4.5	3415	8	US-10-995-561-86	Sequence 86, Appl
c 26	230.5	4.5	126552	12	US-11-121-086-1	Sequence 1, Appli
c 27	230.5	4.5	191684	12	US-11-121-086-2	Sequence 2, Appli
c 28	227.5	4.4	5706	12	US-11-052-554A-519	Sequence 519, App
c 29	227	4.4	3921	12	US-11-052-554A-523	Sequence 523, App
c 30	227	4.4	4983	12	US-11-052-554A-521	Sequence 521, App
c 31	226.5	4.4	38920	7	US-10-330-773-228	Sequence 228, App
c 32	224	4.4	150468	12	US-11-112-908-56	Sequence 56, Appl
c 33	224	4.4	193789	12	US-11-112-908-55	Sequence 55, Appl
c 34	223.5	4.4	2304	12	US-11-052-554A-538	Sequence 538, App
c 35	223	4.3	116856	12	US-11-143-980-1	Sequence 1, Appli
c 36	221.5	4.3	2183	9	US-11-072-512-1064	Sequence 1064, Ap
c 37	221.5	4.3	2736	8	US-10-858-730-38	Sequence 38, Appl
c 38	221	4.3	168516	12	US-11-121-086-3	Sequence 3, Appli
c 39	219	4.3	1566	12	US-11-219-180-1	Sequence 1, Appli
c 40	219	4.3	3711	8	US-10-873-528-321	Sequence 321, App
c 41	218	4.2	11070	12	US-11-075-185-34	Sequence 34, Appl
c 42	218	4.2	78869	12	US-11-075-185-1	Sequence 1, Appli
c 43	217.5	4.2	11070	12	US-11-075-185-34	Sequence 34, Appl
c 44	217.5	4.2	78869	12	US-11-075-185-1	Sequence 1, Appli
c 45	217	4.2	26667	8	US-10-995-561-13375	Sequence 13375, A

ALIGNMENTS

RESULT 1

US-11-052-554A-517
; Sequence 517, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatenIn version 3.3
; SEQ ID NO 517
; LENGTH: 9903
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-517

Alignment Scores:
Pred. No.: 6.27e-05 Length: 9903
Score: 282.50 Matches: 223
Percent Similarity: 37.2% Conservative: 125
Best Local Similarity: 23.8% Mismatches: 365

Query Match:	5.5%	Indels:	226
DB:	12	Gaps:	46
US-09-917-376-1 (1-957) x US-11-052-554A-517 (1-9903)			
Qy 56 AsnValAlaIleGlyGlyGlyGlyPheValAspGlyIleValPheAsnGluGly----			73
Db 7399 AACCGAACATCGCGCGCGCAACATCGGGAC-----TTTAACTCGGGATCGCA 7449			
Qy 74 -----AlaProGlyIleLeuTyValArgThrAspIleGly-----GlyMet 87			
Db 7450 AACACCGGTCGGGGCTAACCGCGGTGTCAAACAATCGGTATCGCAACACCGCAAC 7509			
Qy 88 TyArgTrpAspAlaAlaAsn---GlyArgTrpIleProLeuLeuAspTrpValGlyTrp 106			
Db 7510 TACAACATCGGTGTCGCAACACCGGTAACTACAAC-----ATCGGCTTC 7554			
Qy 107 AsnAsnTrpGlyTyArgGlyVal---ValSerIleAlaAlaAsp----- 120			
Db 7555 GGCAACACCGGCAACAAACAATCGGATCGGCTGTCCGGCGCAACACAGATCGGGTTC 7614			
Qy 121 ---ProIleAsnThrAsnLysValTrpAlaAlaValGlyMetTyThrAsnSerTrpAsp 139			
Db 7615 GGCCCGCTGAACCGCGC-----ATCGCAACAATGGGCTGTTC---AACCTGGGGCAC 7665			
Qy 140 ProAsnAspGlyAlaIleLeuArgSerSerAspGlnGlyAlaThrTrpGlnIleThrPro 159			
Db 7666 AACAACTTTGCC----- 7677			
Qy 160 LeuProPheLysIleuGlyGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaVal 179			
Db 7678 -----ATGGCAACCGCGGCAACTTCAACACAGGCATTCGC 7713			
Qy 180 AspProAsnAsnAspAsnIleLeuTyPheGlyAlaProSerGlyLys---GlyLeuTrp 198			
Db 7714 AACACCGGCAACAAACATCGGTGTTCACACCGGCAACAAACAGTCGGCACTCG 7773			
Qy 199 ArgSerThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThr 218			
Db 7774 CTGACCGGCACGGCTGTTCGGCTTCAGCTCCTGAACTCGCGCGCGCAACACCGGT 7833			
Qy 219 TyrIleAlaAsnProThrAspThrThrGlyTyGlnSerAspIleGlnGlyValValTrp 238			
Db 7834 TTCTTCAACTCCGGCACCGCCAAACACCGC----- 7863			
Qy 239 ValAlaPheAspLysSerSerSerLeuGlyGln---AlaSerLysThrIlePheVal 257			
Db 7864 ---TTGTTCACTCCGGCACCGCAACACCGCTTGTTCACCTCGGCACCGCAACGTC 7920			
Qy 258 GlyValAlaAspProAsnAsnProValPheTrpSerArgAspGlyGlyAlaThrTrpGln 277			
Db 7921 GGCAATCGGCAACATGGCACCGCGCTTCGGCTCGGCTATCCGCGCAGCAGGTG 7980			
Qy 278 AlaValProGlyAlaProThrGlyPheIleProHisLysGlyValPheAspProValAsn 297			
Db 7981 GGATCGCGGCGCAACCACTCGGCAGTTTC---AACATCGGCTGTGTTAACTCGGCACC 8037			
Qy 298 HisValLeuTyIleAlaThrSerAsnThrGlyGlyProTyArgAspGlySerSerGly-As 317			
Db 8038 GGCAATGTCGGCATCGCAACTCGGCACCGCAACGTTCGGCATCGCAACACCGGCACC 8097			
Qy 317 pValTrpLysPheSerValThrSerGlyThr-----Tr 328			
Db 8098 GGCAACACCGGCATCGGAACAGCGCAACTACAACACCGCTTGTCAACGGCGGCTG 8157			
Qy 328 p-----ThrArgIleSerProValProSerThrAspThrAlaAsnAs 342			
Db 8158 GTCAACACCGGCATCGCAACCGCGGCAACCAACACCGGCTGTGTTCAACTCGGCACC 8217			
Qy 342 pTyPheGlyTySerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAl 362			
Db 8218 TTCAACACCGGCATCGCAACCGCGGCACCTACAACACCGGCTCTTACAACACCGGTAGC 8277			

QY	362	aThrGlnIleSerTrp---TrpProAspThrIleile---PheArgSerThr-----	377
DB	8278	TACAAACACGGCATGCGAAACGCGGAGACTTACGGCACCGCGCGCTTCATCAACCGCGCAGC	8337
QY	378	-----AspGlyGlyAlaThrTrpThrArgIleTrpAspTrpThrSer---	391
DB	8338	ATGAACAACGGCTTGCTGGCGCGCGACCGGACGGCGCTGCTGGCGGCCAATACACC	8397
QY	392	-----TyProAsnArgSerLeuArgTyrValLeuAspIleSerAlaG1	406
DB	8398	ATCACCATCGAGCGACCTCGCGGTTCCTCAATGTCGCATCCCGGTCAACATCCCATC	8457
QY	406	uProTrpLeuThrPheGlyValGlnProAsnProValProSerProLysLeuGlyTr	426
DB	8458	ACCG---GGCAGCATCAACCAATGCTCCATCCCGCCATTACGTTCCCCAGA-ATC----	8508
QY	426	pMetAspGluAlaMetAlaIleAsp---ProPheAsnSerAspArgMetLeuTyrGlyTh	445
DB	8509	---GACCGCAGCGGAAGCGTCACATAGGCATCCTCAGTGGCACCGCTTGGCCCGCGT	8564
QY	445	rGly---AlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlyGlnI1	464
DB	8565	CGTCCGATCACCTGCATCGCGGGGAGCGGTGGCGCCCGTGGACACACCCATCGAAT	8624
QY	464	eHisIleAlaProMetValLysGlyLeuGluThrAlaValAsnAspLeuIleSerPr	484
DB	8625	TGACTTCGGCCCG-----TCGCGCGCGATCAACCTCAACATCGGCAA	8666
QY	484	oProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAs	504
DB	8667	GGCCGACGCTCCACCGTGATCAACATCGTGGCGCGCGCGC-----	8709
QY	504	pValThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSerVa	524
DB	8710	-----GCCGCGCCGATC-----AGCAT	8726
QY	524	lAspTyrAlaGluLeuAsnProSerIleValArgAlaGlySerPheAspProSerSe	544
DB	8727	TCCGATCATCGACTTCGGCGCAGCG-----CCCGGCTTCCTCAACGCCACCCAC	8774
QY	544	rGlnProAsnAspArgHisValAlaPheSerThrAspGlyLysAsnTrpPheGlnG1	564
DB	8775	CGGCCCG-----TCGTGGGCTTCCTCAACTGGGGTGCTCG	8810
QY	564	ySerGluProGlyGlyValThrThrGlyGlyThr-----ValAlaAl	578
DB	8811	CAGCGCATCGGCTTGCTGAACCTCGGCACAACTCGGCCCTCTACAACTCGCCACTAG	8870
QY	578	aSerAlaAspGlySerArgPheVal-----TrpAlaProG1	590
DB	8871	CAGCATGGGAAATTCGGGCTTCAAAACATATGGGTGCTGCAGTCGGGTGGCGG----	8925
QY	590	yAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSer-----Tr	606
DB	8926	-----AATTGGGCAACGATCTCGGGCATCTA	8954
QY	606	pAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLy	626
DB	8955	CAACACCGGCTTGGGAGCAGCCGCAATGTC-----TCGGGCTTGCTCAACATCGG	9005
QY	626	sThrPheTyrAla-----LeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyVa	644
DB	9006	CACCAACCTGGCTGGGTGGTTGCAGAACGCGCCG-----	9039
QY	644	lThrPheGlnProValAlaAlaGlyLeu-ProSerSerGlyAlaValGlyValMetPhe-	663
DB	9040	-ACCGAGACGACCTTCAGGTGGCTTGGCCAACTCGGGTCTCT---GGAATCTGGGTAG	9095
QY	664	-----HisAlaValProGlyLysGluGlyAspLeuTrpLeuA	676
DB	9096	CGCAAAACATCGCAACTACAACTGGCGCAGCCCAACATCGCGCTCTCAACCTGGCGAG	9155
QY	676	la-----AlaSerSerGlyLeuTyrHisSerThrAsnGlyLysSerTrpSerA	693

6969 CAGCGAATAGACCGGTTTATCTGCTGAGCGTCACTCCGGTCGCATCGGCCCAACGTCGTA 7028
Db
277 nAlaValProGlyAlaProThrGlyPheIleProHisLysGly---ValPheAspProVa 296
Qy
7029 CTCGGTC-----GGCAGAAATCTACAACGCCGGTGACCTGTCGTCGGCGG 7073
Db
296 lAsnHisValLeuTyrlleAlaThrSerAsnThrGlyGlyProTyraSpGlySerGl 316
Qy
7074 CAATATCATCTCGGAATCCGAGCGTCAGCACACCGCGCGGTGATGCC----- 7125
Db
316 yAspValTrpLysPheSerValThrSerGlyThrTrpThrArgIleSerProValProSe 336
Qy
7126 -----GTCTTCAATCGGCATCTCAGCGCTTCAAC-----ACCCCGC 7163
Db
336 rThrAspThrAlaAsnAspTyraPheGlyTyraSerGlyLeuThrIleAspArgGln---Hi 355
Qy
7164 GCTCAATAT-TCACAGATCCGTTGGTTTCCAGGTGCCGGCAGCATCGACGCGATCA 7222
Db
355 sProAsnThrIleMetValAlaThrGlnIleSerTrpTrpProAspThrIleIlePheAr 375
Qy
7223 CCCT----- 7226
Db
375 gSerThrAspGlyGlyAlaThrTrpThrArgIleTrpAspTrpThrSerTyraProAsnAr 395
Qy
7227 -----GTTCCCGGTGGTCTGACGTTCCCGGC-- 7253
Db
395 gSerLeuArgTyraValLeuAspIleSerAlaGluProTrpLeuThrPheGlyValGlnPr 415
Qy
7254 -----GAATCGTGTCTGAACC-TGGATGTGTCTGTCGGCACCCCC 7293
Db
415 oAsnProProValProSer-----ProLysLeuGlyTrpMetAspGluAlaMetAl 432
Qy
7294 GCGCGCAACCAT-CCGCGCATCACTTCCCGGAGATC----- 7329
Db
432 alieAspProPheAsnSerAspArgMetLeuTyra----- 443
Qy
7330 -----CCGCGGAACCCGAGCGCAACTCTAGTCTATCGCGCGGACATCCCGCTGAT 7382
Db
443 ----- 443
Qy
7383 CAACATCCCGCCACCCCGGCGATTGGGAACACCAACACCGCTTCGTCGCGGCTTCTT 7442
Db
444 -----GlyThrGlyAlaThrLeuTyraAlaThrAsnAspLeuThrLy 457
Qy
7443 CAACACCGCGCGCGGGGCTCGGGTTCGGCAACTTCGGCGCAACATGTCGGGGTG 7502
Db
457 sTrpAsp-----SerGly-----GlyGl 463
Qy
7503 GTGGAACCGAGCGCACACCGCTTTGGCAGCGCGGGTTCGGGTATTGCCAACGTCGGCAC 7562
Db
463 nIleHisIleAlaProMetValLysGlyLeuGluThrAlaValAsnAspLeuIleSe 483
Qy
7563 ACTGCATCGGGTGTCAACCTGCTCGGGCTGTCGGGGATCTACCAACACCGACAC 7622
Db
483 rProProSerGlyAlaPro---LeuIleSerAlaLeuGlyAspLeuGly----- 498
Qy
7623 GCTGCGGCTTTGGAGCGCGGGTGTGGTTCGGGCTTGGCAACGTCGTCATCCTGTC 7682
Db
499 -GlyPheThrHisAlaAspValThrAlaValProSerThrIlePhe-ThrSerProValP 518
Qy
7683 GGGCTTTGTGGCTTCCAACTGGGGGAAAAACCCCATCACCATCGTCAACATCGGGTTGGC 7742
Db
518 heThrThrGlyThrSerValAspTyraAlaGluLeuAsnProSerIleIleValArgAlaG 538
Qy
7743 TAACGTCGGAACCGCAACCTCGGCTTCGGAACATCGGC-AACCTCAACCTCGGTGCGG 7801
Db
538 lySerPheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyG 558
Qy
7802 CCACAT-----GGCAGCGTGAACCTGGGATTCGGCAACATGGCG 7843
Db
558 lyLysAsn-TrpPheGlnGlySerGluProGlyGlyValThrThrGlyGlyThrValala 577
Qy
7844 ACGTGAACCTGGGTTCTCGGCAACATCG---GCGCGGCAACGTCGGGTTTCGGCAATATCG 7900
Db

578 Ala-----SerAlaAspGlySerArgPheValTrpAlaProGlyAspProGlyGln 594
Qy
7901 GCGATGCCAACTTCGGGTTCGGAATTCGG---GTCTGCGCGCGGCTTGGCGGCATCG 7957
Db
595 ProVal-----ValTyraAlaValGlyPheGlyAsnSerTrpAla--- 607
Qy
7958 GCAATATCGGCTGGCAACCCCGCAGCGCAACGTCGTCGGTGGCCCAACATGGGCCCTGG 8017
Db
608 AlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThr 627
Qy
8018 GCAACATCGGGTTGGCAACACCGCACCACTCGGATCGGGTTCACCGCGGACA 8077
Db
628 PheTyraLeuLeuSerAsnGlyThrPheTyraArgSerThrAspGlyGlyValThrPheGln 647
Qy
8078 -----ACCAGTCG 8086
Db
648 ProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHisAlaValPro 667
Qy
8087 GCATCGCGGGTTGAACCTCCG-----GCATCGCAACATTTGGCTTTCACACTCG 8137
Db
668 GlyLysGluGlyAspLeuTrpLeuAlaAspSerGlyLeuTyraHisSerThrAsnGly 687
Qy
8138 GCACAGCA-----ATATCGGCTTCTCACTCCGGAGCTGCCAACTTCGGGT 8185
Db
688 GlySerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnValGlyPheGlyLys 707
Qy
8186 TGTTCAACTCCGCGAGCTTACAACACCGGTATCGCAACTCCG-----GGTGG 8233
Db
708 SerAlaProGlySerSerTyraAlaValPheValValGlyThrIleGlyGly----- 725
Qy
8234 CCAGCACCGGGTTGGTCAACCGCGCGGC-TTCAACACCGGTGTGGCAAAACCGCGGTG 8292
Db
726 ---ValThrGlyAlaTyraArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAsp 744
Qy
8293 TACAACACGGGACGTTCAATGCTGGTGACACCAACACCGG-TGGCTTCAACCCAGGCGAG 8351
Db
745 AspGlnHisGln-TyraGlyAsnTrpGlyGlnAlaIleThrGly-----AspHi 760
Qy
8352 CACCAACACCGGCTGGTTCAACACCGGTAAACCAACACCGGCTGCCCAACCGCGGCA 8411
Db
760 sAlaAsnLeuArgArgValTyrlleGlyThrAsnGlyArgGlyIleVal----- 776
Qy
8412 TGTCAACACCGCGCCCTCATCAGCGGCAACTTTAGCAACGCGCATCTTATGCGGGGCA 8471
Db
777 -TyraGlyAspIleGlyGly-AlaProSerGlySerPro-----SerProSerValS 793
Qy
8472 TTACGAGGGCTTGGCGGCTTCTCTTGGGTACCCCATTCGCTGTTCCCGCGGTGGG 8531
Db
793 erProSerAlaSerProSerLeuSerProSerProSer-----ProSerSerSer---P 810
Qy
8532 CGCCGAGTCAACCGCGACATCGGCCCGCCCAACCATCATTCGCCCAATCCATCCCGTC 8591
Db
810 roSerProSerProSerProSerProSerProSer-----SerSerP 824
Qy
8592 CATCCCGTTGGGCTTCGCGCGATCGGCCACATCGGGCGCGATCAGCATCCCGCAACATCG 8651
Db
824 roSerPro-----S 827
Qy
8652 CATCCCTCGATCCACCTGGGATCGATCCCACTCCACCTTCGACGTCGCGCCCTATCACCGTGA 8711
Db
827 erProSerProSerProSer-----ProSerArgSerProSer- 839
Qy
8712 CCCCATCACCTTCACTCCCTGCGCTTAAAGTTTGGATGCTGCCGCTCTCGGAGATCAGGAT 8771
Db
840 -----ProSerAlaSerProSerProSerProSerProSerProSerProSerS 858
Qy
8772 GACGTCGGAGAGAGCTCCGGATTCAGGTCAAGTCAGCCAGCTTTTCATTCTTCGCGGTGG 8831
Db
858 erSerProSerProThrProSerSerProValSerGlyGlyValLysValGlnTyrl 878
Qy
8832 AC-----CCGACGCGATGTCGCG-----GGGCGAGGTCCTCCATCT 8867
Db

QY 878 yAsnAsnAspSerAlaPro-GlyAspAsnGlnIleLysProGly-LeuGln----- 894
 Db 8868 TCACCACT-----TCACCGTGGGACCATCACTGAACCGGACGACACTGCATCTCC 8921
 QY 895 -----ValValAsnThrGlySerSerValAspLeuSerThrValThrValArgTy 912
 Db 8922 CGGATTCCACCATCCACCGGACCATCCACATCGGCTGGCGTGTGCTGACCATCC 8981
 QY 912 rTrpPheThrArgAspGlyGlySerThrLeuVal-TyrAsnCysAspTrpAlaAlaI 932
 Db 8982 GGGCTTCACCATCCCGGGCGC-----ACCTGATTCCTCCCACTCCCGCTGGGCTCGG 9035
 QY 932 le-----GlyCysGlyAsnIleA 938
 Db 9036 TTTGTTCGGGGCGGACCCACCTTTGATCTCCGACGCTGTTATCGACCGGATCC--C 9092
 QY 938 rGAlaSerPheGlySerValAsnProAlaThrProThrAlaAspThrTyr 954
 Db 9093 GGTGGAGTTACCGGACGACCATCGGCCCCGTGACGCTCCCGATTT 9142

RESULT 3
 US-11-052-554A-525
 ; Sequence 525, Application US/11052554A
 ; Publication No. US20050288866A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sachdeva, et al.
 ; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
 ; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
 ; FILE REFERENCE: 30853/40359A
 ; CURRENT APPLICATION NUMBER: US/11/052,554A
 ; PRIOR FILING DATE: 2005-02-07
 ; PRIOR APPLICATION NUMBER: US 60/589,227
 ; PRIOR FILING DATE: 2004-07-20
 ; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
 ; PRIOR FILING DATE: 2004-02-06
 ; NUMBER OF SEQ ID NOS: 763
 ; SOFTWARE: Patent in version 3.3
 ; SEQ ID NO 525
 ; LENGTH: 11151
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium tuberculosis H37Rv
 US-11-052-554A-525

Alignment Scores:
 Pred. No.: 0.000201 Length: 11151
 Score: 273.00 Matches: 279
 Percent Similarity: 35.2% Conservative: 115
 Best Local Similarity: 24.9% Mismatches: 385
 Query Match: 5.3% Indels: 349
 DB: 12 Gaps: 65

US-09-917-376-1 (1-957) x US-11-052-554A-525 (1-11151)

QY 79 TyrValArg---ThrAspIleGlyGlyMetTyrArgTrpAspAlaAlaAsnGlyArgTrp 97
 Db 2399 TGGGTGGCGGCAACATCGGTGGGCGACACCTGGGCTGG---GCAATCTCGGG----- 2449
 QY 98 IleProLeuLeuAspTrpValGlyTrp-----AsnAsnTrpGlyTyr--- 111
 Db 2450 -----ACGGCAACCTGG---GGTTGGGCAACATCGGCGGCACTGGGGTTTGCCA 2500
 QY 112 -----AsnGlyValValSerIleAlaAlaAspProIle 122
 Db 2501 ACTTGGGCTTGACGCGCGGCGGGGTGGGCAATGTGTTTGGCAATCGCGGCA 2560
 QY 123 AsnThrAsnLys-----ValTrpAlaAlaValGlyMetTyrThr----- 135
 Db 2561 TCACAACTATGCTTGGCGAAACATGGGTGGGCAATATGTTGGTTTGGCCACACCGGCA 2620
 QY 136 -----AsnSerTrpAsp-----ProAsnAspGlyAla---Ile 145
 Db 2621 CGGGCAACATCGGGATCGGGCTGTGTCGGGACCATCGGACCGGGATCGGGGCTTGAAC 2680

QY 146 LeuArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGly 165
 Db 2691 CGGCATCGGCAATATCGGTGTGTTCA-----ACTCCGGCACCGGCAACGTCCGGT 2731
 QY 166 GlyAsnMetProGly-----ArgGlyMetGlyGluArgLeuAlaValAspProAsn 182
 Db 2732 TCTTCAATTCGGGACCGGCAACTTCGCATCGGAACTCCGGCGCTTCAACACCGGGA 2791
 QY 183 AsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGlyLeuTyrArgSerThrAsp 202
 Db 2792 TCGGT-AAT-----AGCGGAACCGGCGGCGGCTCTTC--AATCGCGG 2835
 QY 203 SerGlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsn 222
 Db 2836 AGCTTCAGCAGCGGATCGGCAACACTGCTACTACACACACGCGGAGCTTCAACGCGCG 2895
 QY 223 ProThrAspThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAsp 242
 Db 2896 GACACCAACCGGTGGCTTCAACCCGGGCGGATCAACACCGGCTGG-----TTCAAC 2949
 QY 243 LysSerSerSerLeuGlyGlnAlaSerLysThrIlePhe-----ValGly 258
 Db 2950 ACCGGGATGCAACACCGGTTGGCCAAACGGGCGGACCTTCGGACCCGGCGCTTCATG 3009
 QY 259 ValAlaAspProAsnAsnProValPheTrpSerArgAspGlyGly----- 273
 Db 3010 ACGGCGACTACAGCAACCGGCTGTGTGGCG-----GGCGGTACGAGGCGCTGGTC 3063
 QY 274 -----AlaThrTrpGlnAlaValProGly 281
 Db 3064 GCGCTCCGCGTGGGCGGCCCGATCTCCCAATTCGCGTCACCGTCGACGATCGCGGG 3123
 QY 282 ---AlaProThrGlyPheIleProHisGlyValPheAspProValAsnHisValLeu 300
 Db 3124 GTGGCGCGTGCATGTGGCGGCC-----GTCCGGTATCCCGCGGCGGACGTC--- 3171
 QY 301 TyrIleAlaThrSerAsnThrGly---GlyProTyrAspGlySerSerGlyAspValTrp 319
 Db 3172 GAGATCACCGAGCGGCGCTCGGCTGGTTCGGTTCACCGTCCCGGACCGATAGCATCCC 3231
 QY 320 LysPheSerValThrSerGlyThrTrpThr-----ArgIleSerPro 333
 Db 3232 TCACCTCCCATCGGCGGATCACCGGAAGCGTGGACCTGGCGGCAACACCATCTCGCG 3291
 QY 334 ValProSerThrAspThrAlaAsnAspTyrPheGly----- 345
 Db 3292 ATTCGCGCTCTTGACCGCGTTCGATAGGGCTTTTTCGAGCGGTTCCGCGCTC 3351
 QY 345 ----- 345
 Db 3352 AGTGACCACTTATCACTGATGCGTTCCAAGTTGTTCGGGTGCTTGTTCCTAGAG 3411
 QY 346 -----TyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAla 362
 Db 3412 AACATCATGTGGCGGCTCACGGTTAGCGGTAG-----ATAATTGGTCAACCCG 3462
 QY 363 ThrGlnIleSerTrpTrpProAspThrIleIlePheArgSerThrAspGlyAlaThr 382
 Db 3463 ACACCAATT-----CCCTAACCTCACTTTGGACACACCC----- 3498
 QY 383 TrpThrArgIleTrpAspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAsp 402
 Db 3499 -----CGTGAGCGCTTTTCCGGAATGGT-----TTCACC 3528
 QY 403 IleSerAlaGluProTrpLeuThrPheGlyValGlnProAsnProValProSerPro 422
 Db 3529 ATTCGCGGCAACCCCGTACCGGTGGTATAGGAG---TCGCCAACGAGGGTTCACCT 3586
 QY 423 LysLeu-GlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLe 442
 Db 3587 TCTTCCCGGGTGGGTGACCTTTCCGCGGCGCTCCGC-----CGGGGTCAACGGACTGT 3640
 QY 442 uTyrGlyThrGlyAlaThr-----LeuTyrAlaThrAs 453

Db 3641 CGTGGGGTGGACGGGTTACCGCTGTTGCCGACGGGTTACCCCTCGACACCGTGGCCGG 3700
Qy 453 nAspLeuThrLysTrpAsp-----SerG1 461
Db 3701 CGACCTTCGACGGCACCATCTCATCGGGGATATCCGATCCCGATCATCGATGTGCCGG 3760
Qy 461 yGly-----GlnIleHisIleAlaProMetValLysGlyLeuGluThrAl 477
Db 3761 CGGTGCGGGGTTCCGGCAACACACCGACCGCCCATCGTCCGGGTTCTTCAACACCGGGC 3820
Qy 477 aValasn-----AspLeuIleSerProPr 485
Db 3821 GCGCGGTGGATCGGGGTTCCCAACGTCGCGCGCGGACGTCGGGCTGGTGAACCA-- 3878
Qy 485 oSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGly-----PheThrHisAl 503
Db 3879 ----GGGGACGACGTGTAGCAGG---GGCGGGCTCGGGAGTTGCCAATGCCGGCACGC 3931
Qy 503 aAspValThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSe 523
Db 3932 TGAGTCGG-GCGTGTGAACGTGCGGCTCGGGGATCTCCGGGTGGTACAAACACGACGCC 3990
Qy 523 rValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspProSe 543
Db 3991 CTGGGAGCGGCA-----CCCCGGGGTGGTCTCGGGCATCGGCAACCTCGGCCAG 4041
Qy 543 xSer-----GlnProAsnAspArgHisValAlaPheSerThrAs 556
Db 4042 CAGCTGTCGGGTTCTTGGCAATGGACCGTCTCAACCGGAGCCCATTTGCAATATC 4101
Qy 556 pGlyLysAsnTrpPheGlnGlySerGluProGlyGlyValThrThrGlyGlyThrVa 576
Db 4102 GGGTGGCGCATGTGGCGGT---TCAACACCGGGTGGGCAATGGGGGACCTCAAC 4158
Qy 576 lAlaAlaSerAlaAspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProVa 596
Db 4159 TGGGGTGGCGCAACATCGCGCGGAGAACCTGGGCTG-GGCAATCTCGGACGCGG-- 4215
Qy 596 lValTyrAlaValGlyPheGlyAsn-SerTrpAlaAlaSerGlnGlyValProAlaAsnA 616
Db 4216 -----AACGTGGGTTCCGGCAACATCGGTGCGGCAAGTCGGGTTCCGCAACT-- 4264
Qy 616 laGlnIleArgSerAspArgValasn----- 624
Db 4265 -----CGGTCGGCGGTGGGCTGGCGGCTGGGCAACGTGGGTT 4307
Qy 625 -----ProLysThrPheTyrAlaLeuSerAsn- 633
Db 4308 GAGCAATGCGGCGACCAACATGGGGGCTGGCCAACTGGGGTGGGGCAACATCGGGTT 4367
Qy 634 --GlyThrPheTyrArgSerThrAspGly-----GlyValThrPheGlnProVala 650
Db 4368 GGCCACACCGGACGGGCAACATCGGGATCGGGCTGGTGGCGGATACACAGACCGGCAT 4427
Qy 650 laAlaGlyLeuProSer---SerGlyAlaValGlyValMetPheHisAlaValProGlyL 669
Db 4428 CGGC-GGCTCAACTCGGTAGTGGCAATATCGA---TTGTTCAATTCGGGACCGGCA 4483
Qy 669 yeGluGlyAspLeuTrpLeuAlaAlaSerSer---GlyLeuTyrHisSer----- 684
Db 4484 ATGTCGGGTTCTCAACACCGGACCGGCACTTCGGAGCTGTTCAACTCGGTAGTTTCA 4543
Qy 685 --ThrAsnGlyLysSerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnValG 704
Db 4544 ACACCGGATCGGTAAATAGCGAACCAGGAGTACTGGGCTTCTCAATGCCGCAATTTCA 4603
Qy 704 lPheGlyLysSerAlaProGlySerSerTyrProAlaValPheValValGlyThrIleG 724
Db 4604 ACACCGGATCCCAACCGCGGTCTGTACAAACGCGGACGCTTCAATGTGGGTATACCA 4663
Qy 724 lGlyValThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsnA 744

Db 4664 AC-----ACCGGTGGTTTCAACCGGCGGACATCAACACCGGTGG-----TTCAACA 4711
Qy 744 sAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuA 764
Db 4712 CCGG-CATTATGATACGG-----CACCCGCAACACC 4743
Qy 764 rArgValTyrIle-GlyThrAsnGlyArgGlyIleValTyr---GlyAsp----- 779
Db 4744 GCGCCCTCATGTCCGGGACCGACAGCAACGCGATCTGTGGCGGCGCACACGAGGGC 4803
Qy 780 lIleGlyGly-AlaProSerGlySerProSerValSerProSerValSerProSe 799
Db 4804 CTGTTGGGCTGTCTATGATCATCATGATCCCGCAAT---TCCGATCCGATCACACG 4860
Qy 799 rLeuSerProSer---ProSerProSerSerPro----- 810
Db 4861 ACTGGCGTATCGGCCCATCTCATCCCGACACCATCATCTCCGCGGTGCACCTG 4920
Qy 811 -----SerProSerProSerProSerProSe 821
Db 4921 CAGATCACCGGCGAGCGGACTACAGCTTCAACGTCGCGGACATCCCATCCCGCATC 4980
Qy 821 rSerSerProSerProSerProSerProSerProSerProSerProSe 841
Db 4981 CACATCGGATCAATGGCGTCGTACCGTCGGCTTTCACCGCCCGGAGCCCA---CCCTG 5037
Qy 841 rAlaSerPro-----SerProSerProSerProSerProSe 855
Db 5038 CTGTCCGCCCTGAAGATACCGTAGTTCATCAGCTTCGGCCCATCAG-CTCTCGAA 5096
Qy 855 rProSerSerProSerProSerProSerProSerProSerProSerProSe 875
Db 5097 TATCGATATTCGCGCCCATGGATTTCAAGCTTAGGCGCTGCCGTTCTTGGTCTATCACGG 5156
Qy 875 lGlnTyrLysAsnAsnAspSer-----AlaProG1 885
Db 5157 CCAACT-CGACCAATTCATCTTGAGCCAAATCGTGTGGCGGATCGGTGTGCCCTGG 5215
Qy 885 yAspAsnGlnIleLysProGlyLeuGlnVal----- 896
Db 5216 AGATCGAGCCCATCCCTGGATCGGATTCGTTGAGTGAGTCGATTCCTATCCGATAC 5275
Qy 897 -----AsnThrGlySer----- 900
Db 5276 CTGTTGATATTCGCGCTCGTCATCGATGGGATTCATGTCGGAAGTGGTGGCGATCG 5335
Qy 901 --SerSerValAspLeuSerThrValThrVala-gTyrTrpPheThrArgAspGlyGlys 920
Db 5336 ATGCGTCGTTGGACATCCGGCGGTACGATC-----A 5368
Qy 920 eSerThrLeuValTyrAsnCysAspTrpAlaAlaIleGlyCysGly---AsnIleArgA 939
Db 5369 CAGCACCAACCATT-----TCCGCGATCCCGTGGGCTTCGACATTCGCA 5413
Qy 939 laSerPheGlySerValasn-----ProAlaThrPro 949
Db 5414 CCAGTCCGAGACCCCTCAACATCCGATCCGATCCGCGCGGCGCG 5463
RESULT 4
US-11-052-554A-518
; Sequence 518, Application US/11052554A
; Publication No. US20050288666A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT FILING DATE: 2005-02-07
; PRIOR FILING DATE: 2004-07-20
; PRIOR FILING DATE: 2004-02-06

; NUMBER OF SEQ ID NOS: 763

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 518

; LENGTH: 6615

; TYPE: DNA

; ORGANISM: Mycobacterium tuberculosis H37Rv

US-11-052-554A-518

Alignment Scores:

Pred. No.: 0.000183 Length: 6615
 Score: 270.50 Matches: 263
 Percent Similarity: 32.4% Conservatively: 105
 Best Local Similarity: 23.1% Mismatches: 363
 Query Match: 5.3% Indels: 412
 DB: 12 Gaps: 56

US-09-917-376-1 (1-957) x US-11-052-554A-518 (1-6615)

Qy 53 ThrTrpSerAsnValAlaIleGlyGlyGlyPheValaspGlyIleValPheAsnGlu 72
 Db 5 ACCTGGTCTCC-ACAACGTCGGGAATGTCGGGCTTCCTC-----AACGTC 48
 Qy 73 GlyAla-----ProGlyIleLeuTy-ValArgThrAspIleGlyMetTyArg-Tr 90
 Db 49 GCGCGCTGGATCGGCTGGTGGCGAATGTGGCAACACCACTCTCGGGTATCTACAACGTG 108
 Qy 90 pAspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpG1 110
 Db 109 GGCACGTCGGACCTCTCGACCGCCGCGTTAACTCGG-----GGTGGCAA- 154
 Qy 110 yTyAsnGlyValValSerIleAlaAlaAspProIle-----AsnThrAs 125
 Db 155 -----ATATCGGAACCAATATGTCGGGCTCTGCGCGACGCGC 192
 Qy 125 nLysVal-----TrpAlaAlaValGlyMetTyTrpThrAsnSerTrpAspPr 140
 Db 193 GCGGGTACTCGGCTATTAACTTGGCTTGGCCAAACACCGCAACCTCAACGTGGCTTC 252
 Qy 140 oAsnAspGlyAlaIleLeuArgSerAspGlnGlyAlaThrTrpGlnIleThrProLe 160
 Db 253 GCAAGTCTCGGCGCTTAACTTCG----- 277
 Qy 160 uProPheLysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaValas 180
 Db 278 -----GGGCGGCC 285
 Qy 180 pProAsnAspAsnIleLeuTyPheGlyAlaPro---SerGlyLysGlyLeuTrpAr 199
 Db 286 ACCATCGCCACACAAACGTCGGCATCGGGAACACCGGAATCTTCGATGTCGGCTGGCG 345
 Qy 199 gSer-----ThrAspSerGlyAlaThrTrpSerGlnMe 210
 Db 346 AACCTGGGACGTACAACTCGGCTTCGGAACCTTCGGCAGCAACACCTGGGCTTCGGC 405
 Qy 210 tThrAsnPheProAspValGlyThrTyIleAlaAsnProThrAspThrThrGlyTyGr1 230
 Db 406 AACTCG-----CCAGCTACAACTCGGCTTCGGCAACGTCGCAACG----- 448
 Qy 230 nSerAspIleGlnGlyValTrpValAlaPheAspLysSer----- 245
 Db 449 -----ACAATCTGGTTCGCTAAACGCGGCGCGCAACATCGGCTTT 492
 Qy 246 -----SerSerLeuGlyGlnAlaSerLysThrIlePheValG1 258
 Db 493 GCGAACACCGCGCAGCAAAATGTGGCTTTGGGAACACCGGCGCAACAAATGTCGGCATC 552
 Qy 258 yValAlaAspProAsnAsnProValPheTrpSerArgAspGlyGlyAla----- 274
 Db 553 GGGCTCACGGCAACGGAC-----AGATCGGGTTCGGCAGCTTCAACTCG 597
 Qy 275 -----ThrTrpGlnAlaValPr 280
 Db 598 GGCAGCGGAACATCGGCTTTCAACTCGGGCAGCAACACATCGGATCTTCAATTCC 657

Qy 280 oGlyAlaProThrGlyPheIleProHisLysGlyValPheAspProValAsnHisValle 300
 Db 658 GGCAGCGCAACTTCGGC-ATCGAAACTCGGGCAGCTTCAAC----- 699
 Qy 300 uTyIleAlaThrSerAsnThrGlyGlyProTyArgGly-----SerSerGlyAspVa 318
 Db 700 -----ACTGGCATCGGAACACCGGCAACACCAATACCGGCCTATTCACTCGGCAC-- 753
 Qy 318 lTrpLysPheSerValThrSerGlyThrTrp-ThrArgIle-----SerProValProS 336
 Db 754 -----GTCAACACGGGCGCTTCAACCGGCGCAGCTTCAACACCGGTAGCTT 800
 Qy 336 erThrAspThrAlaAsnAspTyPheGlyTySerGlyLeuThrIleAspArgGlnHisP 356
 Db 801 CAACACCGGCGCTTCAACACCGGTGGCT--TCAATCGGGCAATACCAACACCGGCTA 857
 Qy 356 roAsnThrIleMetValAlaThrGlnIleSerTrpTrpProAspThrIleIle-----P 374
 Db 858 CCTCAACATTTGGCAACTACAACACCGGCATCGCCAAACCGCGCAGTTTGACACCGGGC 917
 Qy 374 heArgSerThrAspGlyGlyAlaThrTrp-----ThrArgIleTrpA 388
 Db 918 TTTCAATACCGGAACACTACAGCAACGGTGTCTTAAGCGCGGATTACCAGGCGCTGG- 976
 Qy 388 spTrpThrSerTyProAsnArgSerLeuArgTyValLeuAspIleSerAlaGluProT 408
 Db 977 -----TCGGCCTCAACCT 989
 Qy 408 rpLeuThrPheGlyValGlnPro-----A 416
 Db 990 GGTGATCATGATGCTCTCCCAATAGCGTCGGCGTGAATATTCCCATCGATATCCCGAT 1049
 Qy 416 snProProValProSerProLysLeuGlyTrp----- 426
 Db 1050 CACCGCTCGCGCGCAACATCACCCTTATGGCGTCACGATTCCGCCCCACCGCGCATAT 1109
 Qy 427 -----MetAspGluAlaMetAlaIleAspPropheAsnSerA 439
 Db 1110 CGTCCTTTCGTCATAGCGGCGCACGACCCACCTTTGGCCCATACCATTCGGAACAT 1169
 Qy 439 spArgMetLeuTy-ArgThrGlyAlaThrLeuTyArgAlaThrAsnAspLeuThrLysTr 458
 Db 1170 CACGGTTCGCGCCCAACGACGACAGTCGCCATAGGAGGGCGCAATACCGCATACCAT 1229
 Qy 458 pAspSerGlyGlyGlnIleHisIle-----AlaProMetValLysGlyLeuG1 474
 Db 1230 AACTGGCGGTGGCGCATTAGGATCCCGTCATCATGATATCCCGCGCGCGCAGTTTCGG 1289
 Qy 474 uGluThrAlaValAsn----- 479
 Db 1290 AAACCTCGAACCAACCCGCTGCTCAAGTTTCTTCAATACCGGCGCGCGCGCTTCGGG 1349
 Qy 480 -----AspLeuIleSerProSerG1 487
 Db 1350 CTTTCGGCAACTTCGCGCGCGCAATTCGGGCTTTTGGAACTTGGCTTCGCGACCTCGG 1409
 Qy 487 yAlaProLeuIleSerAlaLeuGlyAspLeuGly----- 498
 Db 1410 GCGCTCGGGGCTCTCAACGTCGGCGCTGGATCAGGTCTGGCGAACGTCGGGCAACCCAC 1469
 Qy 499 -----GlyPheThrHisAlaAspValThrAlaValProSerThrIlePheThrSerPr 516
 Db 1470 CGTCTCGGGGTTCTACAACACCGACGACCTCGGACCTCGGACGCGCGGCTTCAATTGAG 1529
 Qy 516 oValPheThrThrGlyThrSerVal----- 524
 Db 1530 CCTGGCGCAACATCAGCACCATGATCCCGGCTGCTGCGCGACGACGCGGACCATGTT 1589
 Qy 525 -----AspTyAlaGluLeuAsnProSerIleIleValArgAlaG1 538
 Db 1590 CCTCAACCTGGCTTGGCAACACCGACCCCTCAACGTCGCGCAT----- 1635


```
Db 80 TGGTGGCG---ACGGCGGAGCGCATCGCGGGTTCCTCCGGAGGACCGGGGCTGGGACC 136
Qy 387 TrpAspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGlu 406
Db 137 TGGCGGGGCTCTCGACTCCGACCGCGACG----- 166
Qy 407 ProTrpLeuThrPheGlyValGlnProAsnPro-----ProValProSerPro 422
Db 167 CCGTGG-----GCAAGTCCTATGTGCGGAGGGCGGTTTCCTCACCG 208
Qy 423 LysLeuGlyTyrMetAsp-----GluAlaMetAlaIle 433
Db 209 AC-GGGGGCGGATTCGACCGCGCATCTTCGGCATCTCGCCCGTGGAGCGCTGGCGGATG 267
Qy 434 AspProPheAsnSerAspArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsn 453
Db 268 GACCCG-----CAGCAGCGGTGTCTGTGGAGACCGCG----- 300
Qy 454 AspLeuThrLysTrpAspSerGlyGlyGlnIleHisIleAlaPro---MetValLysGly 472
Db 301 -----TGGGAGACCTTCAGAAATCGCGGAATCGACCCCGGGTTCGTGCACCGC 348
Qy 473 LeuGluGluThrAlaValAsnAspLeuIleSerProProSerGlyAlaProLeuIleSer 492
Db 349 ACCGACGTGCGTGTTCAGCGGAGTGTATGACACAGATTACGGGGCCGACCGCGGGAGC 408
Qy 493 AlaLeuGlyAspLeuGlyPheThrHisAlaAspValThrAlaValProSerThrIle 512
Db 409 GCGGGGAGGCGCTGGAGGG-----CATCTGGCGTGGGACGCGCGGGAGGTGCTGTC 462
Qy 513 -----PheThrSerProValPheThrGlyThr--- 522
Db 463 TCCGGGGGGTGGCTACGCGCTGGCGCTGACCGGGCCCGCGGTGACCGTGACACCGCC 522
Qy 523 -----SerValAspTyrAlaGluLeuAsnProSer 532
Db 523 TGCTGCTCTCCCTGGTAGCGCTGCACCTGGCGGTTACGGCGGTGCGCACGCGGGAATGC 582
Qy 533 IleIleValArgAlaGlySerPheAspProSerSerGlnProAsn----- 547
Db 533 TCGTGGGGCTCGCGGGGTGTCGGGTGATGAGAGCGCGCTGCAAGTCTTCGCGGAGGGCGCCAC 642
Qy 548 AspArgHisValAlaPheSerThrAspGly----- 557
Db 643 TCCCGCAGCGTGGCTCGCCCCCGACGCGCGCTGCAAGTCTTCGCGGAGGGCGCCAC 702
Qy 558 GlyLysAsnTrpPheGln----- 563
Db 703 GGCACCAACTGCTCCGAGGGTGTGGGTGTGTTGCTGGAGCGGCTGTCCGATGCCCGC 762
Qy 564 -----GlySerGluProGlyGlyValThrThrGlyThrValAlaAlaSerAlaAsp 581
Db 763 CGCAATGGGCATAGGTGCTCGCGCTGCTCGTGGC-----ACGCGCGTAACACGAGAC 816
Qy 582 Gly---SerArgPheValTrpAlaProGlyAspProGlyGln----- 594
Db 817 GCGCCAGCAACGCGCTGACCGCGCCCAACGCGCGCTCCAGGAACGCGGTGATCGGCGAC 876
Qy 595 -----ProValValTyrAlaValGlyPhe 602
Db 877 GCGTGGCGAACCGCGGCTGACGTGGCGCGATGTGGAGCGGTGCGAGGCCACCGGCACG 936
Qy 603 GlyAsnSer----- 605
Db 937 GGCACGAGTCTCGCGACCGCGATCGAGGCCAGGCACTCTCTGGCCACCTACCGGCGAGGAG 996
Qy 606 -----Trp----- 606
Db 997 CGGCGGAGGATCAGCCGCTGTGGGTGCTGGGTGCTTGAAGTCGAACATCGGGCATGGCGAG 1056
Qy 607 -----AlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsn 624
Db 1057 GCGGGGCGGCGCGCGCGGTGTATCAAGATGTCTCAGGCCATCGCGCACGGGCTACTG 1116
```

```
Qy 625 ProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyVal 644
Db 1117 CCAAAACACCTCCAGCGCCAGGACCCAGCACGAGTCTGACTCGAGCTCAGGTGCGGTG 1176
Qy 645 -----ThrPheGlnProValAlaAlaGlyLeu 653
Db 1177 TCGTACTGTTCGAGGCGCGGCTCGCGGAGACCGGACACCCCGCGCGCGGGAATC 1236
Qy 654 ProSerSerGlyAlaValGlyValMetPheHisAlaValProGlyLysGluGlyAspLeu 673
Db 1237 TCCTCTCTCGGCGTACGCGGAGGAACGACACGCTGGT---CTTGGAAACAGCG--- 1286
Qy 674 TrpLeuAlaAlaSerSerGlyLeuTyrHisSerThr-----AsnGlyGlySerSerTrp 691
Db 1287 -----ACCCCTGGAAGCGCTGCACCCGAAACACAGCGGAGGACGCGGGCGCTCTCGG 1340
Qy 692 SerAlaIleThrGlyValSerSerAlaValAsnValGlyPheGlyLysSerAlaProGly 711
Db 1341 GCTCGTGGCCACGCGGCGGTAGTCCGTG----- 1370
Qy 712 SerSerTyrProAlaValPheValGlyThrIleGlyGlyValThrGlyAlaTyrArg 731
Db 1371 -----GGTCTGTCCGC 1382
Qy 732 SerAspAspCysGlyThrThrTrpValLeuIleAsnAspAspGln-His---GlnTyrGln 750
Db 1383 CAAGACTCTCGGCGCTCGCGGCTCAGCAGAGCGCTCTGGTCAGGCATCTGGAGTCCGG 1442
Qy 750 yAsn-----TrpGlyGlnAlaIleThrGlyAspHisAlaAs 762
Db 1443 GAGCAGACCAACCCGCTGATGTGGCTGGTGGTGGCCACCACCACCGCGGCGGTGGA 1502
Qy 762 nLeuArgArgValTyrIleGlyThrAsnGlyArgGly----- 774
Db 1503 GCACCGCGCGGTCTCTCGCGCAGCATGTCGAAGGAGGCGATGGCGACGCGCGGGCTCT 1562
Qy 775 -----IleValTyrGlyAspIleGly----- 781
Db 1563 GCGGAGGGGCGGCTGACCCGCTCTGTGTACCGGACAGACCGGACAGACCGGCAAAAC 1622
Qy 782 -----GlyAlaPro-SerGlyS 787
Db 1623 CGTGTTCATCTTCCCGGCAAGCGCCCAATGGTGGGATGGGAGCCCACTCTCTCAA 1682
Qy 787 erProSerProSerValSerPro----- 794
Db 1683 CACCTCACCGGTCTTCGCGCGCGCTGTGAGTGGCGGATGCTCTAGCGCGGTATAC 1742
Qy 795 -----SerAlaAs 797
Db 1743 CGACTGTGCTCATCGAGCTCATCCCGGACGCCCGAGCTCCCTCGCTTGGCGGTGT 1802
Qy 797 erProSerLeuSerProSerProSerProSerProSerProSerPro----- 812
Db 1803 CGACGTGTAGACGCGCGCACCTTCGCGCTGTGCTCTCCCTCGCGCACTCTTGGCAATC 1862
Qy 813 -----SerProSerProSerSer----- 818
Db 1863 GGTGGGATCCACCCCGACGCGCTCATCGGCCACTCCCAAGGCGAAATCGCCGCGCGCTG 1922
Qy 819 --SerProSerSerProSerProSerPro-----SerProSer-----P 832
Db 1923 GGTCCCGGACACCTTCACCTCACCAACCGCCGCAAAATCGTCACCTTCGCGACCGACAC 1982
Qy 832 roSerProSerArgSer-----ProSerP 840
Db 1983 CATGCCCCACCACTCGCGGACAGCGGCGCATGTCTCTCGCCACCCCGCGGACAC 2042
Qy 840 roSerAlaSerPro----- 844
Db 2043 CATCGACCTCACCACTGGCAGCGGCAAACTCTGTGATCGCGGCACACAGACGCCCAACGC 2102
```


Db 5599 GCGCGGATGGTCTGGGTGGTGAATCAGCGCTCGTCCGGTGTCTGGGTGGGGTTCG 5658
Qy 448 ThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlnIleHisAla 467
Db 5659 GGTGTTTCAACGCGCGGCACGCTG-----CATTCGGGTGTGCTGAATTCGCG 5706
Qy 468 ProMetValLysGlyLeuGluGluThAlaValAsnAspLeuLysSerProSerGly 487
Db 5707 TCTGCGATGTGGGGCTGTTCAACACCGCGTGTGGGGTTG-----GGT 5751
Qy 488 AlaPro--LeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThr 506
Db 5752 GCGCGCGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 5790
Qy 507 AlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSerValAspTyr 526
Db 5791 -----CAGTTGTCGGGATGTGGCGGCGGCGGCGG----- 5823
Qy 527 AlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspProSerSerGlnPro 546
Db 5824 -----CTGCATCAGGTGTGGTCTCAATTCGGGTGGCGGATGTGGGTGG----- 5871
Qy 547 AsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnThrPheGlnGlySerGlu 566
Db 5872 -----GCAATGTGCGGTGGGCAATGTGCGGGACTTTAACTGGGTGGCGGCAACGTC 5925
Qy 567 ProGlyGlyValThrThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheVal 586
Db 5926 GGTGGTTCATATGGGTGGGTGGGCAATATCGCGGCAACATGCGGTGGGCAATGTC 5985
Qy 587 ---TrpAlaPro-----GlyAsp-----ProGlyGlnProValValTyr 598
Db 5986 GGTGGGCGCAATTTGGGCTCGGGAATTCGGGGTTAAGCGCGGTCTCATGGGTTCGGT 6045
Qy 599 AlaValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIle 618
Db 6046 AATATCGGTGTGTAATCGCGGAGCTACAAATTCGGGTGG-----GCCAATATG----- 6096
Qy 619 ArgSerAspArgValAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPhe----- 636
Db 6097 -----GGTGTGGGCAATATTTGGGTTCGCTAACACCGCGAGTGGGAATTCGGTATT 6147
Qy 637 -----TyrArgSerThrAspGlyGlyVal 644
Db 6148 GGTGTACCGGTGATATCTACCGGTTCGGTGGTTCATATCCGCGAGCGGAATGTG 6207
Qy 645 ThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyVal---MetPhe 663
Db 6208 -----GGGTGTGTTAATTCGGGACCGGTGAATGTGGGTTCCTTT 6246
Qy 664 HisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHis 683
Db 6247 AACTCTGCACCGGGAAC-----TGGGGGGTGTTCATATTCGGGGAGTTAT----- 6291
Qy 684 SerThrAsnGlyLysSerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnVal 703
Db 6292 AACACCGGATCGGTAATTCGGGATTCGACGCGGGTGTTCACGCGGGTGGGTTC 6351
Qy 704 GlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheValValGlyThrIle 723
Db 6352 AACACGGGTGGTCAATTCGGGTAGCTACACACCGCGCATTTCAACCGCGGCGGCGGCC 6411
Qy 724 GlyGlyValThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsn 743
Db 6412 AAT-----ACGCGCGGTTCACCCGCGGAGTGTCAACACGGGTGGTTCAACACCGGT 6465
Qy 744 AspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeu 763
Db 6466 GACATC-----AACACCGGGTGGCCAACTCCGCGGAC----- 6498
Qy 764 ArgArgValTyrIleGlyThrAsnGlyArgGlyIleValTyrGlyAspIleGlyGly-Al 783

Db 6499 -----GTCAACACCGCGGCTTTCATCTCCGGCACTACAGCAACGGC 6540
Qy 783 aProSerGlySerProSerPro-----SerValSerProSerAlaSerProSerLe 800
Db 6541 GCTTCTCGCGGGGCGACTACACAGGCGCTGCTCGGCTTCTCTACACGACCATCAT 6600
Qy 800 uSerProSerProSerProSerProSerProSerProSerProSerProSerSerPr 820
Db 6601 CCCGAATTCACGTGTCGGAACATCCACGCGTCCGCGGCGCGGACCCATCATCTG 6660
Qy 820 oSerSerProSerProSerProSerProSerProSerProSerProSerProSer 839
Db 6661 TCGATCCAAATTCGCGCAATTCCTTGGACCTACGCGCAACCGGCACATCGGCGGCTTC 6720
Qy 839 rProSer---AlaSerProSerProSerProSerProSerProSerProSerProSer 858
Db 6721 ACCATCCCGCGGTGTCATTTCCCGATCAGGTTCCGATCGACCCAGTCTTCGACCTC 6780
Qy 858 rSerProSerProThrProSerSerProValSerGlyGlyValLysValGlnTyrLy 878
Db 6781 GCGCCCATCAGCGTCCAGGACATCA-----CGATTCCC 6813
Qy 878 sAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGlyLeuGlnValValaAsnTh 898
Db 6814 GCGCTGGGACTCGACCCCG-----CAACCGGTGTACCGTGGCGCCGATATTTCAGTCA 8867
Qy 898 rGlySerSerSerValAspLeuSerThrValThrValArgTyrTrpPheThrArgAspG1 918
Db 6868 GGTCTCATCTCATTCATTCAGCC-----TTACGCTGCTGG----- 6904
Qy 918 yGlySerSerThrLeuVal 924
Db 6905 -GGTTCATCAAGTTAATG 6922
RESULT 7
US-10-995-561-13368
; Sequence 13368, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13368
; LENGTH: 70513
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13368
Alignment Scores:
Pred. No.: 0.00283 Length: 70513
Score: 261.50 Matches: 232
Percent Similarity: 31.5% Conservative: 91
Best Local Similarity: 22.6% Mismatches: 333
Query Match: 5.1% Indels: 380
DB: 8 Gaps: 48
US-09-917-376-1 (1-957) x US-10-995-561-13368 (1-70513)
Qy 86 GlyMetTyrArgTrpAspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGly 105
Db 20976 GGCCACACGCTGGAGCGCATCACACGGGACGACGCGCA----- 21017
Qy 106 TrpAsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsn 125
Db 21018 -----TCTGCACCGCGGAGTACTC 21038
Qy 126 LysValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIle 145

[illegible]

QY 698 SerSerAlaValAsnValGlyPheGlyLysSer----- 708
 DB 22987 AGCAGAGCTTCAGGTGCTGAGCTCCAGCAGAGCTAACTTGCTCTGGGAAATTAAGTGCCA 23046
 QY 709 -----AlaProGlySerSerTyrProAla 716
 DB 23047 AGTGGGAGCAGCTGCTGACGACGAGGAGTGGAACCCAGGAGGCGCAATG----- 23100
 QY 717 ValPheValValGlyThrIleGlyValThrGlyAlaTyrArgSerAspCysGly 736
 DB 23101 -----GGCTCCAGGAGAGGAGGCGCGGGAGCGGGC 23136
 QY 737 ThrThrTrp-----ValleuIleAsnAspAspGlnHisGlnTyrGlyAsnTrp 752
 DB 23137 TTGCTTGGCTTGCGAGGAGACATATGCTCCCTGACTCACCAGCAGCTGGT----- 23190
 QY 753 GlyGlnAlaIleThrGlyAspHisAlaAsnLeuArg---ArgValTyrIleGlyThrAsn 771
 DB 23191 -----GTGACTGAGCTCAGCTCTCCCTGCGAGGTGCAAGAGGGGTGACT 23235
 QY 772 Gly-----ArgGlyIleValTyrGlyAspIle----- 780
 DB 23236 GGAATCTTTCATGGAACACAGCCTCTGTTACCGAGGACCTGTGTACCCACAGTGTCCC 23295
 QY 780 ----- 780
 DB 23296 AAAACATTAGCTGCTCCCTGTCACCTTTTACCAGGACCCACCTCTCTCCC 23355
 QY 781 -----GlyGlyAlaProSerGly----- 786
 DB 23356 AGGGAAGGGGCTCATGTATCATGACGACCAATTTTCTGCTCTTTTACCTCCACCATCAC 23415
 QY 787 ---SerProSerProSerProSerProSerProSerProSerProSerProSerProSer 803
 DB 23416 CTTACTCCCAACCAACCATCACCCCACTACGACCCACCTCCACCATCATCACCCCA 23475
 QY 804 ProSerProSerProSerProSerProSerProSerProSerProSerProSerProSer 822
 DB 23476 CCATCAGCG---TCATCAGCCCCCACCATCATCAGGTATCATCCCCCACCATCATCA 23532
 QY 823 SerProSerProSerProSerProSerProSerProSerProSerProSerProSerAl 842
 DB 23533 CCCCACCTCCACCGTCATCATCCCCCACCATCACCA-----CCCCCACCATCAC 23583
 QY 842 aSerProSerProSerProSerProSerProSerProSerProSerProSerProSerPr 862
 DB 23584 ATCACCATCACCATCACCATCACCATCACCATCACCATCACCATCACCATCACCATCCC 23643
 QY 862 oThrProSerSerPro 868
 DB 23644 ATCCCATCCCATCCCT 23662

RESULT 8
 US-11-052-554A-522/c
 ; Sequence 522, Application US/11052554A
 ; Publication No. US2005028866A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sachdeva, et al.
 ; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
 ; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
 ; FILE REFERENCE: 30853/40359A
 ; CURRENT APPLICATION NUMBER: US/11/052,554A
 ; PRIOR FILING DATE: 2005-02-07
 ; PRIOR APPLICATION NUMBER: US 60/589,227
 ; PRIOR FILING DATE: 2004-07-20
 ; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
 ; PRIOR FILING DATE: 2004-02-06
 ; NUMBER OF SEQ ID NOS: 763
 ; SOFTWARE: Patent in version 3.3
 ; SEQ ID NO 522
 ; LENGTH: 4146
 ; TYPE: DNA

ORGANISM: Mycobacterium tuberculosis H37Rv
 US-11-052-554A-522

Alignment Scores:
 Pred. No.: 0.00043 Length: 4146
 Score: 260.00 Matches: 220
 Percent Similarity: 32.3% Conservative: 99
 Best Local Similarity: 22.3% Mismatches: 384
 Query Match: 5.1% Indels: 287
 DB: 12 Gaps: 43

US-09-917-376-1 (1-957) x US-11-052-554A-522 (1-4146)

QY 20 LeuAlaAlaThrAlaSerPheAlaValAlaAlaLeuGly----- 33
 DB 3315 CTGGCGCGCTGCGCACCTTGGCCATTCGCGGCTTGGCCACCCCGTGTGCGCGT 3256
 QY 34 -----ValleuProIleAlaIleThrAlaSerProAlaHisAlaAlaThr 49
 DB 3255 TGCCGAGCGTTCGCCCATGTTCCGCGTTCGCCGCGGCGCCGAGATGTGCGCAG 3196
 QY 50 GlnProTyrThrTyrSerAsnValAlaIleGlyGlyGlyPheValAspGlyIleVal 69
 DB 3195 TCCGCTTCTCC-----GCCCTTTCGCGCGGTGCTGTGCTGCTGCTGCTGCT 3148
 QY 70 PheAsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyMetTyrArg 89
 DB 3147 GGTGATCGGTGCGCGCGC----- 3127
 QY 90 TrpAspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTyrAsnAsnTrp 109
 DB 3126 -----GCCACCGCTCACCGCTTGTGGACTTGGATGATGG---CGTTGGCGGGCGG 3076
 QY 110 GlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAla 129
 DB 3075 GGCACCATGCGGTGCGCGGAGTGCGCGGCTTCCGTTTACCCTGCGGTGGGTGCGCGC 3016
 QY 130 AlaValGly-----MetTyrThrAsnSerTrpAspProAsnAsp 142
 DB 3015 GCTGCGGCGAGTGGCGGTGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2956
 QY 143 ---GlyAlaIleLeuArgSerSerAspGlnGlyAlaThrTrpGlnIleThr-ProLeuPr 161
 DB 2955 GTGCGCGCGTTCACCGCTTACCGCGGTGCGCGGTGCGCGGTGCGCGGTGCGCGGT 2896
 QY 161 oPheLysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGlyLeuAlaValAspPr 181
 DB 2895 CTGG-----CTGCCA-----GCGTTACCGCC 2875
 QY 181 oAsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerTh 201
 DB 2874 G-----GTAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2831
 QY 201 rAspSerGlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAl 221
 DB 2830 TGCTTCGCGTGCACC-----GTTGCCACCGCGCGCGCGCGCGCGCGCGCGCGCG 2780
 QY 221 aAsnProThrAspThrThrGlyTyrGlnSerAspIleGlnGlyValValTrpVal----- 239
 DB 2779 TGGTGGCGGAGACGAACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2720
 QY 240 -----AlaPheAspLysSerSerSerSerLe 248
 DB 2719 GGTTCGCGCGGTGCGCGCGGTGCGCGCGGTGCGCGCGGTGCGCGCGGTGCGCGCGGT 2660
 QY 248 uGlyGlnAlaSerLysThrIlePheValGlyValAlaAspProAsnAsnProValPheTr 268
 DB 2659 CGCCGCGACCGCGCTACCGCGGTGCGCGGTGCGCGGTGCGCGGTGCGCGGTGCGCGGT 2600
 QY 268 pSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPheLePr 288
 DB 2599 CGCCGCTGTGTGATTGCCACCGAGCGCGCGGTGCGCGGTGCGCGGTGCGCGGTGCGCGGT 2548

Qy 288 oHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsnThrGl 308
Db 2547 -----GCCGCGGT-----GCCCGGTGGCGCGTCGC 2519
Qy 308 yGly-----ProTyrAspGlySerSerGlyAspValTrpLysPheSerValTh 324
Db 2518 CGGGTTGCCCGCGTTCGCGGACGGTATGCTGGTGGCCCTGTGGCCG---GATGCCG 2462
Qy 324 rSerGlyThrTrpThrArgIleSerProValProSerThrAspThrAlaAsnAspTyrPh 344
Db 2461 CGGCGCCACCTTCGCGCGGTGTCCGCGGACCTCC----- 2425
Qy 344 eGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGl 364
Db 2424 -GGTGCCTCCGGCGCCGCCACCGATCGGGGACCGGATCCGCGGTGCGCGCATACCGC 2366
Qy 364 nIleSerTrpTrp----- 369
Db 2365 CTTGGCGCGCTTTCGCGCGTCTCCGCGGTCTCCGCGGTGCGCGCGGTGCGCGCGCTTCC 2306
Qy 369 oAspThrIleIle-----PheArgSerThrAspGlyGlyAlaThr-TripThrA 385
Db 2305 CGAGCGCGTCCGCGGATCAGCGGTTCGCGCGCACCGCGCGCGCCCTGTGGCGG 2246
Qy 385 rGleTrpAspTrpThrSerTyrProAsnArgSerIleuArgTyrValIleAspIleSerA 405
Db 2245 CATTCGCGCGCTGCGCGCGCGCGCGCG----- 2217
Qy 405 laGluProTrpLeuThrPheGlyValGlnProAsnProProVal-----ProS 421
Db 2216 -----GCATTGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2177
Qy 421 erProLysLeuGlyTrpMet-----AspGluAlaMetA 432
Db 2176 CTCACCTTTGGCGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2118
Qy 432 laIleAspProPheAsnSerAspArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAlat 452
Db 2117 CCGCGCGCGCGCGCAATGCTCCGCGCACCGCGGTACCGCATTACACCGCTCGCGCG 2058
Qy 452 hrAsnAspLeuThrLysTrpAspSerGlyGlnIleHisIleAlaProMetValLysG 472
Db 2057 GTTCCGCGCATT-----GGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2007
Qy 472 lyLeuGluGluThr-----AlaValAsnAspLeuIleSerProSerGlyAlaProL 490
Db 2006 CCGCGCGTGAAGCGCGTTCGCGGTCCACCCCAAGTTGCGGTGTCGCGGTGCCAGCC 1947
Qy 490 euleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThrAlaValProS 510
Db 1946 TTGCGCGCGCGCGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1900
Qy 510 erThrIlePheThrSerProValPheThrThrGlyThrSerValAspTyrAlaGluLeuA 530
Db 1899 -----CCAGTGTGCTGCTGCGCGGTGCGCGGTGCGCGGTGCGCGGTGCGCGGTGCG 1851
Qy 530 snProSerIleIleValArgAlaGlySerPheAspProSerSerGlnProAsnAspArgH 550
Db 1850 CCG 1822
Qy 550 isValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGlyV 570
Db 1821 -----GCCTTCGCCACCGAC-----TGGGCGCATTTGCTGTTGTTGCGCGCA 1779
Qy 570 alThrThrGlyThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaProG 590
Db 1778 TTGCGCGCGGTGCGCGGTGCGCGGTGCGCGGTGCGCGGTGCGCGGTGCGCGGTGCGCG 1729
Qy 590 lyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAla--S 609
Db 1728 -----GCCGTTGCGCGGTGCGCGGTGCGCGGTGCGCGGTGCGCGGTGCGCGGTGCG 1686
Qy 609 erGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThrPheT 629

Db 1685 CGCGCGCGCGCGCGCGCG----- 1669
Qy 629 yrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThrPheGlnProV 649
Db 1668 -----GGCGCGCGCGCGCGCGCG----- 1650
Qy 649 alAlaAlaGlyLeuPro---SerSerGlyAlaValGlyValMetPheHisAlaValProG 668
Db 1649 GCCCGCGCGGTTCG 1554
Qy 668 lyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsnGlyG 688
Db 1608 -----GAGGGTGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1554
Qy 688 lySerSerTrpSerAlaIleThr-----GlyValSerSerAlaValAsnValGlyP 705
Db 1553 GTCCGCGCGTTCG 1494
Qy 705 heGlyLysSerAlaProGlySerSerTyrProAlaValPhe-----ValV 720
Db 1493 CG 1434
Qy 720 alGlyThrIleGlyGlyValThrGlyAlaTyrArgSerAspAspCysGlyThrTrp- 739
Db 1433 CCGGTGCTCGCGCGGTTCGCGCGCGCG-----GCCGCGCACCGCGCGCGCGCGCG 1380
Qy 740 -----ValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaI 756
Db 1379 CGCGCGGTGCGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
Qy 756 leThrGlyAspHisAlaAsnLeuArgArgValTyrIleGlyThrAsnGlyArg----- 773
Db 1319 TTACGCGCTTTCG 1260
Qy 774 -----GlyIleValTyrGlyAspIleGlyGlyAlaProSerGly-SerProSer 789
Db 1259 CGCGCGTCAAGCGCGGTTCGCGCGGTACCGCGGTGCGCGGTGCGCGGTGCGCGGTGCG 1200
Qy 790 ProSerVal-----SerProSerAlaSerProSerLeu 800
Db 1199 CGCGCGGTAGCGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
Qy 801 SerProSerProSerProSerProSerProSerProSerProSerProSerProSer 815
Db 1139 CG 1080
Qy 816 ProSerSerProSerProSerProSerProSerProSerProSerProSerProSerPro 834
Db 1079 CG 1020
Qy 835 -----SerArgSerProSerProSerProSerProSerProSerProSerProSer 851
Db 1019 TGGCGCGAACACACTGGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
Qy 852 ProSerSerPro-----SerSerSerProSer----- 861
Db 959 CGCGCTTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCG 900
Qy 862 -----ProThrProSerProSerProSerProSerProSerProSerProSerPro 878
Db 899 CTTGACCG 864
Qy 879 AsnAsnAspSerAlaPro 884
Db 863 AACAAACCG 846

RESULT 9

US-11-205-109-1
; Sequence 1, Application US/11205109
; Publication No. US20050287641A1
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris

APPLICANT: Zazopoulos, Emmanuel
APPLICANT: Staffa, Alfredo
TITLE OF INVENTION: GENE CLUSTER FOR RAMOPLANIN BIOSYNTHESIS
FILE REFERENCE: 3002-2US
CURRENT APPLICATION NUMBER: US/11/205,109
CURRENT FILING DATE: 2005-08-17
PRIOR APPLICATION NUMBER: US/09/976,059
PRIOR FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: US 60/239,924
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 88421
TYPE: DNA
ORGANISM: Actinoplanes sp.
FEATURE:
NAME/KEY: misc feature
LOCATION: (2077)..(3078)
OTHER INFORMATION: ORF 1; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (3118)..(4032)
OTHER INFORMATION: ORF 2; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (4038)..(5048)
OTHER INFORMATION: ORF 3; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (6665)..(5814)
OTHER INFORMATION: ORF 4; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (7703)..(6693)
OTHER INFORMATION: ORF 5; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (9464)..(8130)
OTHER INFORMATION: ORF 6; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (9691)..(10761)
OTHER INFORMATION: ORF 7; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (12751)..(10829)
OTHER INFORMATION: ORF 8; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (13617)..(12802)
OTHER INFORMATION: ORF 9; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (15203)..(13614)
OTHER INFORMATION: ORF 10; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (15591)..(15863)
OTHER INFORMATION: ORF 11; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (15880)..(19035)
OTHER INFORMATION: ORF 12; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (19032)..(39713)
OTHER INFORMATION: ORF 13; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (39713)..(65800)
OTHER INFORMATION: ORF 14; positive strandedness
FEATURE:

NAME/KEY: misc feature
LOCATION: (65826)..(66530)
OTHER INFORMATION: ORF 15; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (66546)..(67370)
OTHER INFORMATION: ORF 16; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (67384)..(70059)
OTHER INFORMATION: ORF 17; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (70059)..(70662)
OTHER INFORMATION: ORF 18; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (70659)..(71906)
OTHER INFORMATION: ORF 19; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (73439)..(71964)
OTHER INFORMATION: ORF 20; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (74216)..(73563)
OTHER INFORMATION: ORF 21; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (75424)..(74213)
OTHER INFORMATION: ORF 22; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (75535)..(76464)
OTHER INFORMATION: ORF 23; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (78110)..(76449)
OTHER INFORMATION: ORF 24; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (79864)..(78107)
OTHER INFORMATION: ORF 25; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (81624)..(79861)
OTHER INFORMATION: ORF 26; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (81909)..(81682)
OTHER INFORMATION: ORF 27; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (82346)..(82062)
OTHER INFORMATION: ORF 28; negative strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (82587)..(84446)
OTHER INFORMATION: ORF 29; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (84481)..(85548)
OTHER INFORMATION: ORF 30; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (85556)..(86845)
OTHER INFORMATION: ORF 31; positive strandedness
FEATURE:
NAME/KEY: misc feature
LOCATION: (87372)..(86803)
OTHER INFORMATION: ORF 32; negative strandedness
FEATURE:
NAME/KEY: misc feature

Db . 51263 AATGGCGGCACG--GTCGTGTGCGCGCCGGGGAGCATCGATGCCCGCTCGGC 51319

QY 874 sValGlnTyrLysAsnAspSerAlaPro 884
 Db 924 GTGCGGGGACAGTCTCACCAGGCTGCGCCG 894

RESULT 11
 US-11-121-086-22/c
 ; Sequence 22, Application US/11121086
 ; Publication No. US20050266459A1
 ; GENERAL INFORMATION:
 ; APPLICANT: POULSEN, TIM S.
 ; APPLICANT: NIELSEN, KIRSTEN V.
 ; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
 ; FILE REFERENCE: 09138.6000-00000
 ; CURRENT APPLICATION NUMBER: US/11/121,086
 ; CURRENT FILING DATE: 2005-05-04
 ; PRIOR APPLICATION NUMBER: 60/567,570
 ; PRIOR FILING DATE: 2004-05-04
 ; NUMBER OF SEQ ID NOS: 107
 ; SOFTWARE: Patent in version 3.3
 ; SEQ ID NO 22
 ; LENGTH: 114801
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-11-121-086-22

Alignment Scores:
 Pred. No.: 0.0149 Length: 114801
 Score: 250.00 Matches: 240
 Percent Similarity: 35.1% Conservative: 103
 Best Local Similarity: 24.6% Mismatches: 385
 Query Match: 4.9% Indels: 252
 DB: 12 Gaps: 48

US-09-917-376-1 (1-957) x US-11-121-086-22 (1-114801)

QY 103 TrpValGlyTyrAsnAsnTrpGlyTyrAsnGlyValValSerIle----- 117
 Db 60092 TGGGCGGGTGGGGCTTTGGGAATGAAGAGGAGGCGCAGTCTCTGCTGGGGCGAGGCGAGG 60033
 QY 118 -----AlaAlaAappProIleAsnThrAsnLysValTrpAlaAlaValGlyMetTyrThr 135
 Db 60032 GCAAGTGGAGTGGAGCGCTCCAGCCACCAGCTGTTCAGCGCCTCCAAATTCCTCCATCC 59973
 QY 136 -----AsnSerTrpAspProAsnAspGlyAlaIle-----LeuArg 147
 Db 59972 TCTCTGGCTCAGCTTCCCTTATGCCGATGGGCGCATCTCTGGTAGGAGGCTTGGA 59913
 QY 148 SerSerAspGln-----GlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGly 165
 Db 59912 GGAGCTGATGAAGTGAAGTGAAGTGGCTCTGGGCGGTGAGG---GTGGCTGTGAGGCTAGCT 59856
 QY 166 GlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAappProAsnAsnAspAsn 185
 Db 59855 GCGTCCCTGCC----- 59844
 QY 186 IleLeuTyrPheGlyAlaProSerGly-----LysGlyLeu 197
 Db 59843 ---CTCCCTGCTGGTGGCGGAGGAGTCTGCTGCCATCTCTGAGGCGCTCTCAGGGCATA 59787
 QY 198 TrpArgSerThrAspSerGlyAlaThrTrpSer---GlnMetThrAsnPheProAsp--- 215
 Db 59786 GAGAG-----GGTTCCTCATGGAGTGGGACCTGATTTTTCCTCCAGTGA 59739
 QY 216 ---ValGlyThr-----TyrIleAlaAsnProThrAsp----- 225
 Db 59738 TGTCTGTGCTCTAAATAACATCTCCAGGATTAATGGTCCAGGCCCCCAACAGCACCAGGT 59679
 QY 226 -----ThrThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAsp--- 242
 Db 59678 GATTGAAAAATGGTTTATGGCCCG-----TGGGTGAAATTCAACTTG 59634
 QY 243 -----LysSerSerSerSerLeuGlyGlnAlaSerLysThrIle 255

Db 59633 CATCTGAGGTGATTTTCCAGATGGATTTTCTCTCTGGGAGGCGGCACATTCACACA 59574
 QY 256 PheValGlyValAlaAaspPro-----AsnAsnProVal-PheTrpSerArgAspGlyG 273
 Db 59573 CTAATATAATAGCAGCTCTCGACATTCATCACCCATAGTGGGGGGACACCTGTCTGG 59514
 QY 273 yAlaThrTrpGlnAlaValProGlyAla-ProThrGlyPheIleProHisLysGlyValP 293
 Db 59513 GGCACAGGGGCGAGGTGCTCCCTGGGTAACTCCAGGTCTCTTAATCTCAGCA----- 59461
 QY 293 heAspProValAsnHisValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyrAspG 313
 Db 59460 --GATCTGGGGGGTGGTCTTAGAGATGCCCTTCACGAGTCCAGGCGCCAGGGAGG 59403
 QY 313 LysSerGlyAspValTrpLysPheSerValThrSerGlyThrTrpThr----- 329
 Db 59402 GGGAGCGGGAGCCCCCAGGAAA-----TGAAACAACTCTAGTGGGTGCCCCAGGGA 59349
 QY 330 --ArgIleSerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyrSerGlyL 349
 Db 59348 GGAGGCTGGAAGGAGG-CCCCAGGCTGAG-----GGTAGGACCCGCGC 59308
 QY 349 euThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTrpTrp 369
 Db 59307 TAGTCTTCCCGAGGAGCCACAGGCCCCCTCTGAGCAGACTCCAGGCGGCGCTGGC 59248
 QY 369 roAspThrIleIlePheArgSerThrAspGlyGlyAlaThrTrpThrArgIleTrpAspT 389
 Db 59247 CCATCTTGTGCAAGAGGCTGGAGTCTCTGGGTGAGCTGG----- 59207
 QY 389 rpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGluProTrpL 409
 Db 59206 -----CAGTTGACCTTGGATCTGACTGTCTCCCTGAAA 59173
 QY 409 euThrPheGlyValGlnProAsnProProValPro---SerProLysLeuGlyTrp-Met 427
 Db 59172 TTCCTTTTACCCTCACTCAGGCTCTCGAGTCCCGCAGCAGCGCCAGCAGGATGGCAG 59113
 QY 428 AspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeu----- 442
 Db 59112 GAGGAA-----GCCCTTCTCCGACGGGCCCTCACCGACCTCTGCTGAGGGCGCTTGA 59059
 QY 443 ---TyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGly 461
 Db 59058 GCCCAGGCTCTTGTCTCAGCAGCTTGGCCGTGGAGATTGGGAATAITGGAGTTTGGG 58999
 QY 462 Gly-----GlnIle 464
 Db 58998 GGATCTCGTGGCTGAGGCGCTGTGCTCATGAAGGACTCCTTGAGTTTGAAGATGAGCT 58939
 QY 465 HisIleAla-----ProMetValLysGlyLeuGluGluThrAlaValAlaAsnAsp 480
 Db 58938 CACCTTCTGCACAGCAATAGGTCTATGTGTGGAACACTGACCACCAACCAAC--- 58882
 QY 481 LeuIleSerProSerGlyAlaPro-----LeuIleSerAlaLeuGlyAspLeuGly 498
 Db 58881 ---CTCCATCCCGCGGGGCGACGTGCTTTATCACCCTTCTAGTGTAGTACAGG 58825
 QY 499 GlyPheThrHisAlaAspValThrAlaValProSerThrIle-PheThrSerProValPh 518
 Db 58824 GGCTCAAAAGGGGCGAGGTTCTCAGGTACACAGCAGGTGAGCAGCAGCAGTGTAG 58765
 QY 518 eThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaG 538
 Db 58764 GGCTGGGGGTGCTCT-----GCCAGGGGCTG 58738
 QY 538 ySerPheAspProSerSerGlnProAsnAspArgHisValAla-----PheSe 554
 Db 58737 AGGTGTGCCCATGCAAGAATCCAGGGGTCCGTGGGCCAGGCGCCCTGCCAAGAGTTGGG 58678
 QY 554 rThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGlyValThrThrGlyG 574

Db 58677 GGCTGAGGAGGAAGGGGCAATTCACCCCTGACGCC-----CTGCCGTGGGGGT 58624
Qy 574 yThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaProGlyAspProGlyG 594
Db 58623 AGAAGAGAGATGCCCTCTCTGGGGCCGAGGAGGCTGGGCCCTCCTCCCA 58564
Qy 594 n---ProValValTrpAlaValGly-PheGlyAsnSerTrpAla---AlaSerGlnGlyV 612
Db 58563 CTTGCCACAGGTGAACCTGACAGAGTGGGA-----TGGAGCTGGCATCCCAAGGGA 58510
Qy 612 alProAlaAsnAlaGlnIleArgSerAspArgValAlaAsnProLysTrpPheValAlaLeu- 631
Db 58509 ACCTCTGGAGTATCACTCCACATGCTTGCAGTAATATTGGGCCCTTCTCCCTCC 58450
Qy 632 SerAsnGlyThrPhe-----TyrArgSerThr 640
Db 58449 AGTAATGAACCTTATTACTTTTAACTAAGATTAAATGTTCTCTGTGAGTTCAAGATACT 58390
Qy 641 AspGlyGlyValThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGly 660
Db 58389 GACTGCCCTGTGAGGCTCAGTCCACGCTCTGCT---CTTCTGCTCTGGGGGCCA--- 58337
Qy 661 ValMetPheHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGly 680
Db 58336 -----CCCCCAGGAGGAGGAGCCAGGAGCCCTGGGGAAGCAGAGAAG 58292
Qy 681 LeuTyr---HisSerThrAsnGlyGlySerSerTrpSerAlaIleThrGlyVal----- 697
Db 58291 TTGTTTCTCCAGGAATCTCCAGGAGGAGATCCATTATCATCATCACCAGTGTCCACCATG 58232
Qy 698 SerSerAlaValAsnValGlyPheGlyLysSerAlaProGlySerSerTyrProAlaVal 717
Db 58231 AGCAGGCCCATACACCATCATGCGCAGCTCCACGAGCAGCCCATCACCAGCAGCAGC 58172
Qy 718 PheValValGlyThrIleGlyGlyValThrGlyAlaTyrArgSerAspCysGlyThr 737
Db 58171 GC-ATCATCACCCACCCACCCACCCACCCACCCATCACCAGCAGCAGCAGCAGC 58113
Qy 738 ThrTrpValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThr 757
Db 58112 ACA-----ATCAC-----ACCATCACT 58095
Qy 758 GlyAspHisAlaAsnLeuArgValTyrIleGlyThrAsnGlyArgGlyIleValTyr 777
Db 58094 GTAACCAACTTCCATC-----ACCACCACTTACCTCACCATC 58053
Qy 778 GlyAspIleGlyAlaProSerGlySerProSerProSerValSerProSerAlaSer 797
Db 58052 CCCACCATCACTACCATCAGTACCACCATCACTTATTTTACACCATCACTCA 57993
Qy 798 ProSerLeuSerProSerPro---SerProSerSerProSerProSerProSerPro 816
Db 57992 CCACCCACCATCACTACCATGATACCATGATACCATGATACCATGATACCATGATAC 57933
Qy 817 SerSerSerProSerSerProSerProSerProSerProSerProSerProSerArg 836
Db 57932 CCATCATCACCATCACCATCTCCATCCTCCATCACCATCACCATCACCATCAC 57874
Qy 837 SerProSerProSerAlaSerProSerProSerProSerProSerProSerProSerPro 856
Db 57873 ACACCCACCGCAGTAAACACCATCTCCATCTCCATCTCCATCTCCATCTCCATCTCC 57814
Qy 857 SerSerSerProSerProThrProSerProSerProSerProSerProSerProSerValGln 876
Db 57813 TCATATCATCATCACTACCTCCATCTCCATCTCCATCTCCATCTCCATCTCCATCTCC 57766
Qy 877 TyrLysAsnAsnAsp-----CAC 57766
Db 57765 TATGGTCGTGTAAGTACGAGCCCTCCTACTACTGTCTCATGAGCTGTCTAGTACCATG 57706
Qy 882 ---SerAlaProGlyAspAsnGlnIleLysPheProGlyLeuGlnValValAsnThrGlySer 900
Db 57705 CTGACAAGCCCA-----CAGTGGTGTGGGGCTGCCATGTCTTGGGACTAGTTCC 57655

Qy 901 SerSerValAspLeuSerThrValThrValArgTyrTrpPheThrArgAspGlyGlySer 920
Db 57654 TCAGGCACAGGAGGCTCCACGCTGCTGATTCATTCG-----AGGGGTTC 57607
Qy 921 SerThrLeuValTyrAsnCysAspTrpAlaAlaIleGlyCys-----934
Db 57606 CTGCAAAAGTCTTT-----ATGGGGTGGCATTTAGAGTGGGGAG 57565
Qy 935 GlyAsnIleArgAlaSerPheGlySerValAsnPro 946
Db 57564 GGGGAGGTAGGAGGAGGCTCTGGAGGACTTCATGCCA 57529
RESULT 12
US-11-052-554A-529/c
; Sequence 529, Application US/11052554A
; Publication No. US20050289866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 529
; LENGTH: 3240
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-529
Alignment Scores:
Pred. No.: 0.00149 Length: 3240
Score: 247.50 Matches: 200
Percent Similarity: 31.9% Conservative: 94
Best Local Similarity: 21.7% Mismatches: 380
Query Match: 4.8% Indels: 252
DB: 12 Gaps: 33
US-09-917-376-1 (1-957) x US-11-052-554A-529 (1-3240)
Qy 12 ArgSerArgArgLeu-----ValSerLeuLeuAlaAlaThrAlaSerPheAlaVal 28
Db 2238 CGTGGCCGCGCGCGCGCTTCCCGCGCGCACCGCGCTGCGCGTGGCGTGGT 2179
Qy 29 AlaAlaAlaLeuGlyValLeuProIleAlaIleThrAlaSerProAlaHisAlaAlaThr 48
Db 2178 GCCGCCCATATGCGCGGTGGCGCGCTGCGCGCTGCGCGTGGCGTGGCGT 2119
Qy 49 ThrGlnProTyrThr-----TrpSerAsnValAlaIleGlyGlyGly 63
Db 2118 GCCCGTCCCGCACCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGGC 2059
Qy 64 PheValAspGlyIleValPheAsnGlnGlyAlaProGlyIleLeuTyrValArgThrAsp 83
Db 2058 GCCGTGGCGGTGGCGGTGGCGCGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGC 1999
Qy 84 -----IleGlyGlyMetTyrArgTrpAspAlaAlaAsnGlyArgTrpIlePro 99
Db 1998 GCCCGCGTGGCGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 1939
Qy 100 LeuLeuAspTrpValGlyTyrAsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAla 119
Db 1938 ACCACACACCGTGGCG 1879
Qy 120 AspProIleAsnThrAsnLysValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAsp 139
Db 1878 -----GTTACCGCATTTGCCACCTGTGGCGCGT----- 1849

QY 140 ProAsnAspGlyAlaIleLeuArgSerSerAspGlnGlyAlaThrTrpGlnIleThrPro 159
DB 1848 -----GCCGCGCTGCGCGCGCTGCGCGCTCCACCTCGCCGCGCGCGCC 1798
QY 160 LeuProPheLysLeuGlyAlaAsnMetProGlyArgGlyMetGlyGluArgLeuAlaVal 179
DB 1797 ACCACCGCCAGCGTTACCGCTTGTCCACCGGAGCGCGC---GGGAGCGGTGCCACCGCC 1741
QY 180 AspProAsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArg 199
DB 1740 GAC-----GGTGGGTGCGCGCGCGGTGCGCGCGCGCACC 1708
QY 200 SerThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyr 219
DB 1707 GCGGTGCGCGCGCGCGCACC---CTGACCCCGCGGTCCCGCCACCGGTGCGCGCC 1657
QY 220 IleAla-----AsnProThrAspThr 226
DB 1656 TTGGCGCGCGCGCGCCCTTTCGCGCGCGCGCGCTTACGGCGCGCGCGCGGTGCC 1597
QY 227 ThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAspLysSerSer 246
DB 1596 CGCGGTGCGCGCGGTGATATTCGCGCGT-----GCCGCGCGT 1558
QY 247 SerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspProAsnProVal 266
DB 1557 CCGCGCGCGCGCGCGCACCTTGACACCGCGCGCGC----- 1516
QY 267 PheTrpSerArgAspGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPhe 286
DB 1515 -----GCCGCGCGCGCGCTGCGCACCTGTGCGCGGTGCGCGCGCGCGCGGTGCC 1462
QY 287 IleProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsn 306
DB 1461 GCCGCGCGCTCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1402
QY 307 ThrGlyGlyProTyrAspGlySerSerGlyAspValTrpLysPheSerValThrSerGly 326
DB 1401 CCGCGCGGTGCGCGCGGTGTGTAGCTGAGAGCGTGGCGCGCGCGCGCGCGCGCGCG 1342
QY 327 ThrTrpThrArgIleSerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyr 346
DB 1341 -----ACCGCGTGGCGCGCGGTGCCAGCGGTGCGCGTAGC-----GCCGCG 1300
QY 347 SerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSer 366
DB 1299 ACCG 1249
QY 367 TrpTrpProAspThrIlePheArgSerThrAspGlyGlyAlaThrTrpThrArgIle 386
DB 1248 -----ACCTCTTGGCGCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCG 1220
QY 387 TrpAspTrpThr-----SerTyrProAsnArgSerLeuArgTyrValLeuAsp 402
DB 1219 TGGTGTGGGCG 1161
QY 403 IleSerAlaGluProTrpTrpLeuThrPheGlyValGlnProAsnProValProSerPro 422
DB 1160 TTGCGCGCGCTTGGCG-----CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1119
QY 423 LysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeu 442
DB 1118 CCGCGCGCACCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1063
QY 443 TyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGly 462
DB 1062 ---GGTGGCG 1027
QY 463 GlnIleHisIleAlaProMetValLysGlyLeuGluThrAlaValAsnAspLeuIle 482
DB 1027 ----- 1027

QY 483 SerProProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHis 502
DB 1026 -----CCG 982
QY 503 AlaAspValThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThr 522
DB 981 GCGAGCGGTACCGCGGTGCTGCGCGC-----GGCACCCACCGCGCAC 940
QY 523 SerValAspTyrAlaGluLeuAsnProSerIle----- 533
DB 939 GCGGTTCGCGCTTTCGCGCAATCCACCGCTTTCGCGCGCTATTCCTCGGTACCGCGC 880
QY 534 -----IleValArgAlaGlySerPheAspPro 542
DB 879 GTCTCCG 820
QY 543 SerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPhe 562
DB 819 GCG 790
QY 563 GlnGlySerGluProGlyGlyValThrThrGlyGlyThrValAlaAlaSerAlaAspGly 582
DB 789 -----AGCACCGCGCGGTTCACCGGTGCGCGCGCGCGCGCGCGCGCG 751
QY 583 SerArgPheValTrpAlaProGlyAspProGlyGlnProValValTyrAlaValGly--- 601
DB 750 TCC-----GCCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 718
QY 602 PheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAsp 621
DB 717 GCGGAGATGCG 658
QY 622 ArgValAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAsp 641
DB 657 ACCGTGGCGCGCGCGGT----- 640
QY 642 GlyGlyValThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyVal 661
DB 639 -----GGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 619
QY 662 MetPheHisAlaValProGlyLysGlyAspLeuTrpLeuAlaAlaSerSerGlyLeu 681
DB 618 CCGCGTCTCGCGC-----CTTTCGCGCGGTGCGCGCGCGCGCGCGCG 586
QY 682 TyrHisSerThrAsnGlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlaVal 701
DB 585 GCGCGCGGTGTGCGCGCGGTGCTGCTGTGCGCGCGCGCGCGCGCGCGCGCGCG 526
QY 702 AsnValGlyPheGlyLysSerAlaProGlySerSerTyr-----ProAlaValPheVal 719
DB 525 GCCATTCG 466
QY 720 ValGlyThrIleGlyGlyValThr-----GlyAlaTyrArg 731
DB 465 GCGCGTTCG 406
QY 732 SerAspAspCysGlyThrThrTrpValLeuIleAsnAspAspGlnHisGlnTyrGlyAsn 751
DB 405 CCGGTGTGCG 373
QY 752 TrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgArgValTyrIleGlyThrAsn 771
DB 372 -----GCCGCGCTTACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 328
QY 772 GlyArgGlyIleValTyrGlyAspIleGlyGlyAla-ProSerGlySerProSerProse 791
DB 327 GGAGCGGTACCGCGCTTCCGCGCGGTGCGCGCGGTGCGCGCTTACCGCGCGCGCAC 268
QY 791 rValSerProSerAlaSerProSerLeuSerProSerProSerProSerProSerProse 811
DB 267 GCTGTGCTGCGAGTGACCGTTCGCGCGCGGTGCGCGCTTTCGCGCGCGGTGCGCGCG 208
QY 811 rProSerProSerProSerProSerProSerProSerProSerProSerProSerProse 831

Db 207 GCGCGGTTGCCCGG-----TTGCTTGTGGCGCGGTCCTCCGGTCCGGCGCTCC 154
 Qy 831 rProSerProSerArgSerProSerProSerAlaSerProSerProSerProSerProSe 851
 Db 153 GCGCGTTCCGCGTTCG 94
 Qy 851 rProSer-----SerSerProSerProSerProSerProSerProSerProSerSe 867
 Db 93 GCGCGCGGTTGGT 34
 Qy 867 rProVal 869
 Db 33 GCGCGTTG 27
 RESULT 13
 US-11-136-527-2622
 ; Sequence 2622, Application US/111136527
 ; Publication No. US20050287570A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wyeth
 ; APPLICANT: Mounts, William M
 ; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
 ; FILE REFERENCE: 031896-041000 (AM101086)
 ; CURRENT APPLICATION NUMBER: US/11/136,527
 ; CURRENT FILING DATE: 2005-05-25
 ; PRIOR APPLICATION NUMBER: US 60/574,294
 ; PRIOR FILING DATE: 2005-05-26
 ; NUMBER OF SEQ ID NOS: 362830
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 2622
 ; LENGTH: 7231
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 US-11-136-527-2622

Alignment Scores:
 Pred. No.: 0.00267 Length: 7231
 Score: 247.50 Matches: 235
 Percent Similarity: 32.1% Conservative: 124
 Best Local Similarity: 21.0% Mismatches: 368
 Query Match: 4.8% Indels: 391
 DB: 12 Gaps: 53
 US-09-917-376-1 (1-957) x US-11-136-527-2622 (1-7231)
 Qy 86 GlyMetTyrArg-----TyrAspAlaAlaAsnGlyArgTrpIleProLeu--- 100
 Db 3121 GGAATATACAGAACTTCTGCTTCTGGGACAGAGAAAT---CAGTGGACTTCTCTTGG 3177
 Qy 101 -----LeuAspTrpValGlyTrpAsnAsnTrpGly----- 110
 Db 3178 AGGAGACGATCTAGAACTTCTACCTCTGGAATAGATGAGCGCAGTGCTTCTTACTGG 3237
 Qy 110 ----- 110
 Db 3238 AAGAGGAGTCTAGACAGCTGCTCTCGAGTAGAGGACCTCAGTGGACTTCTTCTTGG 3297
 Qy 111 -----Tyr-AsnGlyValValSerIleAlaAlaAspProLeAs 123
 Db 3298 AGAAGAAGGTTTCAGAAACATCTACCTCTGGAATAGAGGACATTAGTGTACTTCCGACTGG 3357
 Qy 123 nthraenLysValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGI 143
 Db 3358 AGAAAGTCCAGAACTTCTGCTTCTGGAGTGGGAGACTTGAAGTGGACTTCCCTCAGGAGG 3417
 Qy 143 yAlaIleLeuArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPhely 163
 Db 3418 AGAAAGTCTAGAAACATCTGCTTCTCAGTGTAGAG----GATGCACCCAGCTTCTTACTGA 3474
 Qy 163 sLeuGlyGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAs 183
 Db 3475 AAGAGGAGGCTTAGAGACTTCTGCCTCTCGAATAGAGGACATCAGTGTGCTTCTTACTGG 3534

Qy 183 nAspAsnIleLeu-----TyrPheGlyAlaProSerGI 194
 Db 3535 AAGAGAAATCTAGAAACTTCTGCTCTGGAGTAGAGGATGTCAGTGGACTTCTCTCTGG 3594
 Qy 194 yLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPhPr 214
 Db 3595 AAAAGAAGGCTTAGAGACTTCTGCTCTCTGGAATA-----GAGGACATTAGTGTGTCTCC 3648
 Qy 214 oAsp-----ValGlyThrTyrIleAlaAsn---ProTh 224
 Db 3649 TACTGAAGCAGAGGCTCTAGAAACTTCTGCTCTCGGGGATATGTCAGTGGATTCCTTC 3708
 Qy 224 rAspThrThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAspLysSe 244
 Db 3709 TGGAGAAGATGGTACAGAAACCTCTACTTCTGGAGTAGAGGCTGTGAGT-----GG 3759
 Qy 244 rSerSerSerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspProAsnAs 264
 Db 3760 TCTTCTCTTGGAGGAGAGGCTCTAGAAACCTCTGCCTCTGGAGTAGAAGATCTTGGTCT 3819
 Qy 264 nProValPheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProTh 284
 Db 3820 TCTTACAGAGATAGTCTAGAGACTTCTGCTCTCAGGAGTAGATGTTACTGGTATCTCTTC 3879
 Qy 284 rGly-----PheIleProHisLysGlyValPheAspProValAsnHi 298
 Db 3880 TGGAGAGAGGACACAGAAACCTCTGTCTCT-----GGGCTAGGTAGTACCTTAGTGG 3933
 Qy 298 sValLeuTyr-----IleAlaThrSerAsnThrGly----- 308
 Db 3934 ACTTCTCTTGGACAGAGGCTCTAGAGACTCTAGCTTTCAGGAGCTGAGGACCTTGGTGG 3993
 Qy 309 -----GlyProTyrAsp-----GlySerSerGlyAspValTrpLysPhe---- 321
 Db 3994 CTTCCTCTTGGAAAGAGACTTGTAGGCTCTGCTCTCGAGCCCTCGGACTTCTGGCAA 4053
 Qy 322 -----SerValThrSerGlyThrTrpThrArgIleSerProValProSerTh 337
 Db 4054 GCTACCTTCTGGAACTCTAGGAAGTGTGTCAACTCCAGAGACTAGTGGCTTCTCTCTGG 4113
 Qy 337 rAspThrAlaAsnAspTyrPheGlyTyrSerGlyLeu-ThrIleAspArgGlnHisProA 357
 Db 4114 ATTTAGTGTGAA-----TATTCTGGAGTGACATTTGGAAGTGGCCCATCTC 4161
 Qy 357 snThrIleMetVal-----AlaThrGlnIleSer----- 366
 Db 4162 TGGCCTCTGCTGACTTTTAGTGACTTCCATCCGGCTTCCCAACAGCTCTCTCTTGTGGACAG 4221
 Qy 367 -----Trp-----TrpProAsp----- 370
 Db 4222 TACTTAGTGGAGTGATCAGCCACCACCTGCGCAGTGAAGTGAAGAGGGGACCAT 4281
 Qy 371 -----ThrIleIlePheArgSerThrAspGly-----GlyAlaThrTrpThrArgIle 387
 Db 4282 CAGCGTCAGTGGTTCAGGAGAAGAGTCAGGCGCCCTCAGTGTGGACAGTAGTGC 4341
 Qy 387 rAspTrpThrSerTyrProAsnArgSer-----LeuArgTyrValLeuA 402
 Db 4342 G-GACATTAGTGTCTCTCTTTCAGGAACTGAACCTCAGTGGCCAAACATCCGATCTCTTG 4400
 Qy 402 spIleSerAlaGluProTrpLeuThrPheGlyValGlnProAsnPro----- 417
 Db 4401 ATCTCAGTGGAGAAACATCTGGATTTTGTATGTTAGTGGACGCCATTTGGGTCTTCTG 4460
 Qy 418 -----ProValProSerProL 423
 Db 4461 GCAGTGTGAGGAGACATCTGGGATTCCTGAGTGTGAGTGGACAGCGCTCAGGAGTCTCTG 4520
 Qy 423 ysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAsp----- 439
 Db 4521 ACACCACTGAGATATCTGAGCTTAGTGGACTGTCTCTCTGCGCAACACAGATGTCAGTGGAG 4580

QY 440 -----ArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrL 457
Db 4581 AAGGCTGTGGAAATCTCTTTGGCAGTGGCCAACTCCTGGGCAATACATCT----- 4630
QY 457 ysTrpAspSerGlyGlyGlnIleHisIleAlaProMetValIysGlyLeuGluThrA 477
Db 4631 -----GTGAGTGA---GAACCTCTG 4649
QY 477 laValAsnAspLeuIleSerProProSerGlyAlaProLeuIleSerAlaLeu----- 494
Db 4650 GGATTCTGTATCTCAGTGGGCAACCTTCAGGTTCCAGTGTCTCAGCGGAGCAACACACCTG 4709
QY 495 -----GlyAspLeuGlyGlyPheThrH 502
Db 4710 GAACCCCTGACCTGGCTTCTGGCGCATGAGTGGCAGTGGAGATTTCTTCTGGCATACGT 4769
QY 502 isAlaAspValThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyT 522
Db 4770 TTGTGGAC-----ACCAGTCTTATTGAAGTGACCCCTA 4802
QY 522 hrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspP 542
Db 4803 CCACATTAGAGAGAGAG-----GGGTAGGCTCTGGGAAC 4841
QY 542 roSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpP 562
Db 4842 TCAGTGGCTCCTCTCTGGGAGACAGATCTGTCTGGCAGATCTGGGATGTGGAT---G 4898
QY 562 heGlnGlySerGluProGlyValThrThrGlyGlyThrValAlaAlaSerAlaAspG 582
Db 4899 TCAGTGGCAATCTCTGGGCAATCGATTCCAGTGGACTCATATCCCCCACTCCAGAA 4958
QY 582 lySerArgPheValTrpAlaProGlyAspProGlyGlnProValValThrAlaValGlyP 602
Db 4959 TCAGT----- 4963
QY 602 heGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsn---AlaGlnIleArgSera 621
Db 4964 -----GGCTCCCAAGCGGAGTAGCTGAGTCACTGAGG 4997
QY 621 spArgValAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrA 641
Db 4998 AGGTCTCTGGAGTTGAGACTGGGAGCAGCTTGTCTCTGGGAGCATTC-----G 5045
QY 641 spGly----- 642
Db 5046 ATGGAGTGGACTGTCTCAGTGTTCCTACGGTCTCTTTGTAGACAGAACTTTGGTGG 5105
QY 643 --GlyValThrPheGlnProValAla-----AlaGly-----LeuP 654
Db 5106 AGTCTATAACCTGGCTCTACTCTCAAGAGCTGGAGAGGCGCTTCCAGCATTTTGG 5165
QY 654 roSerSerGlyAlaValGlyValMetPheHisAlaValProGlyLysGlyGlyAspLeu- 673
Db 5166 AATTCACTGGTGGCC-----CATTTCTGGAACACACAGACATATCTGGGACCTCT 5213
QY 674 -----TrpLeuAlaLaser----- 678
Db 5214 CTGGGTCTTTAGACAAAGCACATGGCGCCTGGGTGGACAGAGCCAGACAGAGCCAC 5273
QY 679 -----SerGlyLeuTyrHisSerThrAsn-----GlyGlySerSert 691
Db 5274 CAAGTCCCAATATTTAGTGGGACTTCTTAGCACCACTGATGCAAGTGGAGATCCA 5333
QY 691 rpSerAlaIleThrGlyValSerSerAlaValAsnValGlyPheGlyLysSerAlaProG 711
Db 5334 TAACCTGCCCACTGGC-----AGTGGGAA 5360
QY 711 lySerSerTyrProAlaValPheValValGlyThr-----IleGlyGlyValThrGlyA 729
Db 5361 CCTCGGGCTTCCGGAAGTCACTTTAATCACTTCAGAGTTAGTGGAGGCGTGAAGTGAAC 5420
QY 729 laTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAspGlnHisGlnT 749

Db 5421 CAACCTGTTTCCAGGAACCTTGGC----- 5443
QY 749 yrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgArgValTyrIleG 769
Db 5444 -----CATGCTCTTCTATGACA-----TATACCTCCCGCTCTTT---G 5480
QY 769 lyThrAsnGlyArgGlyIleValTyrGlyAspIleGly----- 782
Db 5481 AGCCAGTGGGGAAGCCTCAGCATCTGGGACCTTGGTGACCTGTAACTTTTCCCG 5540
QY 783 -----AlaProSerGlySerProSerProSerValSerProSera 796
Db 5541 GGTCTGGGTAGAAGCTTCAGTCCCAAGAGGAGGAGTACCCTCTTACCTCAGG 5600
QY 796 laSerProSerLeuSerProSerProSerProSerProSerProSerProSerP 816
Db 5601 CTGGAGTGGAGTGTCTGCTGCTGCTGAGCCAGCAGTCAATTGCTGAGTTCAGATC 5660
QY 816 roSerSerSerProSerSer-----ProSerProSerProS 829
Db 5661 TGCATGGAATCACTCTGCTCCCTGAAACAGATCTGGAGATGACAAACCCAGGCACGG 5720
QY 829 erProSerProSerPro-----SerArgSerProSerProSerProS 845
Db 5721 AGGTGAGCAGCAACCCATGGACCTTTTCAGGAAGGACACAGGAGGATCCGCTGCTCCAG 5780
QY 845 erProSerSerProSerProSerProSerProSerProSerProSerProS 865
Db 5781 AAGTCACTGGAGAAATCTAGCATCTCCGACATAGATCGAGCAGCTTCAGGTGTGCTT 5840
QY 865 erSerSerProValSerGlyGlyValIysValGlnTyrIysAsnAspSerAlaProG 885
Db 5841 TTGCCAGCCCATGACTTCTGGAGCAGGACTGAAATCATGTCAGATGCTCTGAT--- 5896
QY 885 lyAspAsnGlnIleLysProGlyLeuGlnValIleAsnThrGlySerSerValAspL 905
Db 5897 -----CATACCTCAGAGTCAATGTTACCG 5921
QY 905 euSerThrValThr-----ValArgTyrTrpPheThrArgAspGlyGlySerThrL 923
Db 5922 TCAGCACCACCTCCAGAGTCCAGTGGGCGCCAGTCTACCCAGCACCTTACAGAGAC 5981
QY 923 euValTyrAsnCysAspTrpAlaAlaIleGlyCysGlyAsnIleArgAlaSerPheGlyS 943
Db 5982 TT-----CAAGAAATCGAT 5996
QY 943 exValAsnProAla-----ThrProThrAlaAspThr 953
Db 5997 CCCCAATCCCTCATACTCAGGAGAGAGACCCCAACAGCAGAAACA 6043
RESULT 14
US-11-056-470-2
; Sequence 2, Application US/11056470
; Publication No. US20050261263A1
; GENERAL INFORMATION:
; APPLICANT: SANTI, Daniel V.
; APPLICANT: MYLES, David C.
; APPLICANT: TIAN, Zong-Qiang
; APPLICANT: HUTCHINSON, C. Richard
; APPLICANT: JOHNSON, Robert G., Jr.
; APPLICANT: ZHOU, Yi-Qing
; APPLICANT: FENG, Li
; TITLE OF INVENTION: BENZOQUINONE ANSAMYCINS
; FILE REFERENCE: 300622007510
; CURRENT APPLICATION NUMBER: US/11/056,470
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 10/212,962
; PRIOR FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 60/310,779
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/389,255
; PRIOR FILING DATE: 2002-06-14

; PRIOR APPLICATION NUMBER: US 60/393,929
 ; PRIOR FILING DATE: 2002-07-03
 ; PRIOR APPLICATION NUMBER: US 60/395,275
 ; PRIOR FILING DATE: 2002-07-12
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 6360
 ; TYPE: DNA
 ; ORGANISM: Streptomyces geldanus
 ; US-11-056-470-2

Alignment Scores:

Pred. No.: 0.00257 Length: 6360
 Score: 247.00 Matches: 227
 Percent Similarity: 32.4% Conservative: 77
 Best Local Similarity: 24.2% Mismatches: 338
 Query Match: 4.8% Indels: 300
 DB: 12 Gaps: 49

US-09-917-376-1 (1-957) x US-11-056-470-2 (1-6360)

Qy 24 AlaSerPheAlaValAlaAlaLeuGlyValLeuProIleAlaIleThrAlaSer--- 42
 Db 3409 GCCTCGTCGGAGGCGGAGCGGTCGCGATCCACCGCGCGTCTGGAGCGCGACTG 3468
 Qy 43 -----ProAlaHisAlaAlaThrThrGlnProTyr 52
 Db 3469 CACGCGGGCAGCTTCTGCTGCCCTCCGACCGCGCGGACAGGTGACCTGCTGCGGTC 3528
 Qy 53 ThrTrpSerAsnValAlaIleGlyGlyGlyPheValAspGlyIleValPheAsnGlu 72
 Db 3529 GCCTGGAACACCGTCTGTCGACCGCGGC----- 3558
 Qy 73 GlyAlaProGlyIle---LeuTyrValArgThrAspIleGlyGlyMetTyrArgTrpAsp 91
 Db 3559 GCGCGTCCCGGTCGCGGTCGAGTCCGC-----CCGTCGCGGAC-----GAC 3603
 Qy 92 AlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyr 111
 Db 3604 GCCTTCGTGATCGC-----CTGACCGCGCTCGGCGCAGCGTGGCTCGGTG 3654
 Qy 112 AsnGlyValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaVal 131
 Db 3655 GACTCGCTACCTTTCGCGGCGGTGGACCGCGCCAGCTCAAGATCGGCACGCGCGAC 3714
 Qy 132 GlyMetTyrThrAsnSerTrpAspProAsn-----AspGlyAlaIle----- 145
 Db 3715 GCCTGTGGACGTCCTCGGACGAGACCTCGCTGCCGCGCGCGGTCTCTCTGGGCC 3774
 Qy 146 ---LeuArgSerSerAspGlnGlyAlaThr-----TrpGlnIleThrProLeuProPhe 162
 Db 3775 CGCTCGCGAGTTCGGCCACCGGGGACCGGGGCTACCGGGCCACAGGGGACGGCGGA 3834
 Qy 163 LysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGlu-----ArgLeuAlaValAsp 180
 Db 3835 GGCCAGGGGCGGCTTCGGA-CGTCCTCGTGGCGGATACGCGCGCTGGGCGGAAGA 3893
 Qy 181 ProAsnAsnAsnIleLeuTyrPheGlyAlaProSerGlyLysGly----- 196
 Db 3894 CCTACCG-GACCGC-----CGACCGCGCGCGCGCGGAGCTCACCGGCC 3937
 Qy 197 ---LeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPheProAsp 215
 Db 3938 GCCTGTGGAGGAGATCCACCGGTGGG-----TCGCGCGACGACCGCATGGCGGGA 3988
 Qy 216 ValGly-----ThyTrilleAlaAsnProThrAspThrGlyTyrGlnSerAsp 232
 Db 3989 CGCGGTCGCGGTGTCACCGCGCGCGGTTCGCGTCCACGACGACACGAGGTACCG 4048
 Qy 233 Ile-----GlnGlyValTrpValAlaPheAspLysSerSerSerLeu 248
 Db 4049 ACCCGCGCGCCACCGGCTCTGGGCGCTGTCGCTCGGCGCCAGGCGGAACACCGCGGGC 4108

Qy 249 Gly-----GlnAlaSerLys-ThrIlePheValGlyValAlaAs 261
 Db 4109 GGGTGGCCCTGCTGGATGCCGACGAGCGTCCGAGAACTGCCCGCGGGTGTGTCCG 4168
 Qy 261 pProAsnAsnProValPheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProG 281
 Db 4169 GGGACGAGCCCAACT-----GGCGTGGCGGTGGCGCGCGGTGTG----- 4208
 Qy 281 yAlaProThrGlyPheIleProHisLysGlyValPheAspProValAsnHisValLeuTy 301
 Db 4209 -----GGTCCACGCTCACCCGCGTCCAGCCCGGCTGCGCGTCCGCGCAGG 4258
 Qy 301 rIleAlaThrSerAsnThrGlyGly---ProTyrAspGlySerSerGlyAspValTrpLy 320
 Db 4259 CGTGTGGCATCTGGACTCGGCGGAGTACGGCACCTCGACAATCTGGCGTGTCTCCCG 4318
 Qy 320 s-----PheSerValThrSerGlyThrTrpThrArgIleSerProVa 334
 Db 4319 ACAGCGCGGACCGCGCACCGCGCGCGGTGCGGA-TCGAGGTCCGCGCGCGC 4377
 Qy 334 lProSerThrAspThrAlaAsnAspTyrPheGlyTyrSerGlyLeuThrIleAspArg 354
 Db 4378 GGGCTCAACT-----TCCGGGATG----- 4396
 Qy 354 nHisProAsnThrIleMetValAlaThrGlnIleSerTrpTrpProAspThrIleIlePh 374
 Db 4397 -----TCTGTGCTCTCGGCATGTATCCG 4422
 Qy 374 eArg-----SerThrAspGlyGlyAlaThrTrp-----Th 384
 Db 4423 GCGCGTGGTATCGGCACGAGCGCGGTGTGTACCGCAAGTCCGTCGCGCGCTC 4482
 Qy 384 rArgIleTrpAspTrpThrSer-----TyrProAsnArgSerLeuArgTy 399
 Db 4483 ACGGCGCTGGCGGTGGCGCACCGGTGTGTCTCTCGGTCTGTTCGACCGC----- 4540
 Qy 399 rValLeuAspIleSerAlaGluProTrpLeuThrPheGlyValGlnProAsnProProVa 419
 Db 4541 -----TGCC-ACCGCGCGCGCGCGCGTATCCGGAT 4574
 Qy 419 lProSerProLysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAs 439
 Db 4575 GCGCGAG-----GGCTGGTGC----- 4590
 Qy 439 pArgMetLeuTyrGlyThrGlyAla-----ThrLeuTyrAl 451
 Db 4591 -----TTCGCGACGCGCGCGGTGCGGTGCGCTATCTGACGCGCGCTGTACGC 4640
 Qy 451 aThrAsnAspLeuThrLysTrpAspSerGlyGlyGln-----IleHisIleAlaProMe 469
 Db 4641 GTTCACGACCTCGGAGGTCCAGCGCGGCGAGAGGTCTGTGTGCGACCGCGCGCGG 4700
 Qy 469 tValLysGlyLeu-GluGluThrAlaValAsnAspLeuIleSerProProSerGlyAlaP 489
 Db 4701 CGGTGTGGCATGCGCGCGGTCCAGCTCGACAGCACCTTCGCGCGCACCGTCTCGGCAC 4760
 Qy 489 rOleuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThrAlaValP 509
 Db 4761 CGCCACCCCTTCCA-----AGCACCGACGACCTCCACCGCGGTGGCGTTC 4805
 Qy 509 roSerThr-----IlePheThrSerProValPheThrThrGlyThrSerValA 525
 Db 4806 CGCGAAGCGGTCCCTCCAGCGCGGACCTCGCTTACCGCGACACTTCCCGCCACCG----- 4861
 Qy 525 spTyrAlaGluLeuAsnProSerIleValArgAlaGlySerPheAspProSerSerG 545
 Db 4862 -----CCGAGTGTCTCACTCCCTCACCAGCGGACGACATCGACGC 4904
 Qy 545 lnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGly- 564
 Db 4905 CTCCTCGGAC-----TTCCTCAACCCCGCGCGCGGTCTCTCTGGAGATGGGAA 4952

QY 565 SerGluProGlyGlyValThrThrGlyGlyThrValAlaAlaSerAla-----AspGly 582
 Db 4953 GACCCACTGCGGAGCGCGGAGGTGCGGCGCGGATCCGAGGTCACTACCGGCC 5012
 QY 583 SerArgPheValTrp---AlaProGlyAspProGlyGlnProValValTyrAlaValGly 601
 Db 5013 GTTCGATCTCGTGGGAGGCGCGCGGAGCGGTGCGGAGTTCGTCACCACTGGT 5072
 QY 602 PheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAsp 621
 Db 5073 -----GGAGCTGTTCCGAGCGCGCGGATCCGAGCC----- 5102
 QY 622 ArgValAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAsp 641
 Db 5103 -----GCTCGCGGTACGCGAGTGGGATCACCACCGCGCGCC 5138
 QY 642 GlyGlyValThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyVal 661
 Db 5139 CGAGCGGTTCGCTGGATGA-GTCAGGGCGGATACCG----- 5176
 QY 662 MetPheHisAlaValProGlyGlyGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeu 681
 Db 5177 -----GCAAGATCGTGTCTCACCC 5194
 QY 682 TyrHisSer-----ThrAsnGlyClySerSerTrpSerAlaIleThrGlyValSerSer 699
 Db 5195 TGCCACGCGCTCGACCGGACCGGACCGCTCTGTGTACCG----- 5236
 QY 700 AlaValAsnValGlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheVal 719
 Db 5237 -----GTGGACGGGACCTTCGGCGCCACGATCG-----CCGCGCACCTTCTCA 5281
 QY 720 ValGlyThrIleGlyGlyValThrGlyAlaTyrArgSerAspCysGlyThrThrTrp 739
 Db 5282 CCCAGCAG----- 5290
 QY 740 ValLeuLeuAsnAspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAsp 759
 Db 5291 -----GCGCAGCCCATCTGCTGTGTACGCGCC-----GGGAC 5326
 QY 760 HisAlaAsnLeuArgArgValTyrIleGlyThrAsnGlyArgGlyIleValTyrGlyAsp 779
 Db 5327 CGGACGACCTG----- 5338
 QY 780 IleGlyAlaProSerGlySerProSerProSerValSerProSerAlaSerProSer 799
 Db 5339 -----GCGCACACACCTGACCA---CGAACTCACCGAATCGCGCCACCGTCC 5386
 QY 800 LeuSerProSerProSerSerSerProSerProSerProSerProSerSerSer 819
 Db 5387 GCATCACCG-----CCTGCGACACCGCGCGGACCACTCGCGCGCTCTCTCG 5437
 QY 820 ProSerSerProSerProSerProSerProSerProSerProSerProSerProSer 839
 Db 5438 CCGACATCCCGCGGAC-----ACCCCTCACCGCGGTGTGTCCACACGCGCGGA 5488
 QY 840 ProSerAlaSer-----ProSerProSerSerSerProSerProSerSerSer 855
 Db 5489 CCCTCGACGAGGTGCTGACCGCGCTCACCGCGGACCGCTCGACACCGTCTTCCGCG 5548
 QY 856 ProSerSerProSerProSerProSerProSerProSerProSerProSerProSer 870
 Db 5549 CCAGGTGCGCGGTGCTGACCATCTCCACGACCTCACCGCGGAC 5593

RESULT 15

US-11-052-554A-541/c
 ; Sequence 541, Application US/11052554A
 ; Publication No. US2005028866A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sachdeva, et al.
 ; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
 ; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
 ; FILE REFERENCE: 30853/40359A

; CURRENT APPLICATION NUMBER: US/11/052,554A
 ; CURRENT FILING DATE: 2005-02-07
 ; PRIOR APPLICATION NUMBER: US 60/589,227
 ; PRIOR FILING DATE: 2004-07-20
 ; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
 ; PRIOR FILING DATE: 2004-02-06
 ; NUMBER OF SEQ ID NOS: 763
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 541
 ; LENGTH: 2352
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium tuberculosis H37Rv
 ; US-11-052-554A-541

Alignment Scores: 0.00197 Length: 2352
 Pred. No.: 243.00 Matches: 186
 Score: 32.2% Conservative: 55
 Percent Similarity: 24.8% Mismatches: 289
 Best Local Similarity: 4.7% Indels: 222
 Query Match: 12 Gaps: 33
 DB: 33

US-09-917-376-1 (1-957) x US-11-052-554A-541 (1-2352)
 QY 251 AlaSerLysThrIlePheValGlyValAlaAspProAsnAsnProValPheTrpSerArg 270
 Db 2175 GCGGCGCCACCGACCCACCGCGGTGCGGACCGCGGACCGCGCC----- 2128
 QY 271 AspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPheIleProHisLys 290
 Db 2127 ---GGCGCGCGCGTGGACACAGCAGCGCGCGGTGCCACCGCC----- 2086
 QY 291 GlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsnThrGlyGlyPro 310
 Db 2086 ----- 2086
 QY 311 TyrAspGlySerSerGlyAspValTrpLysPheSerValThrSerGlyThrTrpArg 330
 Db 2086 ----- 2086
 QY 331 IleSerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyrSerGlyLeuThr 350
 Db 2085 ---CCGCGCGTCCCGACCGAGGAAACCAATCC-----GCCGCGCGCGCGCGCGCC 2035
 QY 351 IleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTrpTrpProAsp 370
 Db 2034 GTGCGGAACAGACGCGC-----TGCGTACCGCGCGCGCGCGCGCGCC 1987
 QY 371 ThrIle---IlePheArgSerThrAspGlyGlyAlaThrTrpThrArgIleTrpAspTrp 389
 Db 1986 GGTGTCCAGGTTTTCGCGGTACCGCGCGTCCGCGCGCGCGCGCGCGCGCGCGCAT 1927
 QY 390 ThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGlu-ProTrpLe 409
 Db 1926 TCCGGCATTCG 1872
 QY 409 uThrPheGlyValGlnProAsnProValProSerProLysLeuGlyTrpMetAspG1 429
 Db 1871 -----CCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1848
 QY 429 uAlaMetAlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThrGlyAlaThrLe 449
 Db 1847 ----GCACCGATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1793
 QY 449 uTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlnIleHisIleAlaProMe 469
 Db 1792 CGTCGCCCAACAAAGCGCGGTTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1736
 QY 469 tValLysGlyLeuGluThrAlaValAsnAspLeuIleSerProSerGlyAlaPr 489
 Db 1735 G-----CCACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1700
 QY 489 cLeuIleSerAlaLeuGlyAspLeuGlyPheThrHisAlaAspValThrAlaValPr 509

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 13:55:49 ; Search time 42.6467 Seconds
(without alignments)
1434.576 Million cell updates/sec

Title: US-09-917-376-3
Perfect score: 4036
Sequence: 1 ATTQPTWNSNAVIGGGFVD.....YIGTNGRGIYDGGAPSG 740

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	219.5	5.4	2736	2	US-09-252-991A-30227
2	181.5	4.5	3892	2	US-09-328-352-5503
3	167.5	4.2	1164	2	US-09-902-540-12627
4	166.5	4.1	1300	2	US-09-902-540-9932
5	162.5	4.0	772	1	US-08-410-784A-5
6	159.5	4.0	1060	2	US-09-248-796A-14123
7	158.5	3.9	774	2	US-09-346-237-8
8	157.5	3.9	688	2	US-09-489-039A-9813
9	154	3.8	1548	2	US-09-252-991A-22301
10	152.5	3.8	1751	2	US-09-136-574A-44
11	152	3.8	3472	2	US-09-408-020-4
12	151.5	3.8	776	2	US-09-346-237-4
13	148.5	3.7	750	6	5457037-3
14	148.5	3.7	751	6	5457037-5
15	148.5	3.7	776	2	US-09-346-237-7
16	147	3.6	3623	2	US-09-341-461-2
17	146	3.6	1187	2	US-09-949-016-6513
18	146	3.6	1749	2	US-09-640-419C-28
19	146	3.6	2169	2	US-09-949-016-6930
20	143	3.5	894	2	US-08-362-525-22
21	143	3.5	894	2	US-08-971-692-15
22	143	3.5	4861	2	US-09-919-497-70
23	142	3.5	893	2	US-09-605-703B-434
24	142	3.5	1216	2	US-09-134-000C-5130
25	141	3.5	720	2	US-09-296-284-25
26	141	3.5	754	2	US-09-296-284-4
27	140.5	3.5	387	2	US-09-252-991A-21962

28	140.5	3.5	1000	2	US-09-352-159-25
29	140.5	3.5	1000	2	US-09-352-168-25
30	140.5	3.5	1000	2	US-09-771-045B-25
31	140.5	3.5	1000	2	US-09-770-564A-25
32	140.5	3.5	1000	2	US-09-658-835C-25
33	140	3.5	880	2	US-09-902-540-10174
34	139.5	3.5	712	2	US-09-248-796A-14274
35	139	3.4	826	2	US-09-328-352-7515
36	137.5	3.4	656	2	US-09-248-796A-14824
37	137.5	3.4	1298	2	US-09-252-991A-30579
38	137.5	3.4	2316	2	US-09-949-002-314
39	137	3.4	483	2	US-09-902-540-12467
40	137	3.4	619	2	US-09-252-991A-26352
41	137	3.4	902	2	US-09-902-540-10065
42	136.5	3.4	522	2	US-09-538-092-1096
43	136.5	3.4	2415	2	US-09-949-002-398
44	136	3.4	2628	1	US-08-570-311-14
45	135.5	3.4	1205	2	US-09-352-159-29

ALIGNMENTS

RESULT 1

US-09-252-991A-30227
; Sequence 30227, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30227
; LENGTH: 2736
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30227

Query Match	5.4%	Score 219.5	DB 2	Length 2736
Best Local Similarity	21.2%	Pred. No. 6.6e-07		
Matches	190	Conservative 94	Mismatches 324	Indels 289
Gaps	48			
QY	59	GMNNWGY-----NGVV--STAADPI-NTNKVWA-----AVGMYTNSWDPNDGAILR-SS	103	
Db	1305	GGGNTFTPATPLANGTIVVNAVAQAPAGTGGGTTTVDVAPNTPVVNPNSGNLNGTA	1364	
QY	104	DQATATQITPLPFKLKGNMPCRMGERLAVDPNNNLIYFGAPSGKGLWRSTDGATWSQ	163	
Db	1365	EPGSIIVLT-----DGN--GNPIQOTTA-----DGSNNSTFTG-----SQ	1398	
QY	164	MTNFPDVGTYI-ANPTDTTGYQSDIQGVVWVAFDKSSSSLSQ---ASKTIFGVGVADPNNP	219	
Db	1399	LPN---GTVNVVNTASDAAGNTS---APATTVTDSLSLPQVDPSPNGSVISTADAGNT	1451	
QY	220	VF-----WSDGGATQAVPGAP-----TGFP	242	
Db	1452	IIITDGNNGPIQGVATDGSNNSTFTGPIPLDPTVVNVVARSPPNVDSAPAVITVDGVAP	1511	
QY	243	HKGVPDPVNVHLYIATNTGGY---DGSNGDVWKFVST-SGTWTRISVPV-----	289	
Db	1512	AAPVIDPSNGTEISGTAEAGATVITDGGNPIGATDGSNNWTFPTGTPLANGTVINA	1571	
QY	290	-STPDANDYFGYSGITIDRQHPNTIMVATQISWPPDTTIFRSTDGATWTRIDWTSYPN	348	
Db	1572	VAQDPAGNTSGPASVTVDAIAP---PAPVNPNSGVISGTAAGAT-VILTDDGNP-1625		

Query Match	4.2%; Score 167.5; DB 2; Length 1164;
Best Local Similarity	19.8%; Pred. No. 0.0013;
Matches	173; Conservative 89; Mismatches 333; Indels 279; Gaps 42;
Qy 6	YTNWVAIGGGFVDGIVFNEGACILYVRTDIGMYRWDAANGRWIPLL--DWVGNW 63
Db 179	WAWGRNAAGQLG--DGTTLDRAREVAVPEGLTSVTAV----AAGGSHVALLRSDGSVM-AW 231
Qy 64	GYN-----GVYSIAADPINTNKVMAAVGMWYNS-----WDPNDGAILRSSD 104
Db 232	GYNALGQLGDGTTVDRLTPVRVSLGAVVAAGSYFSMALQSDGTVTWTCGEPGEGQLD 291
Qy 105	QGATWQITPLPFLKLGNNMPCRMGERLAVDPNNNIIYFGAPSGKGLWRSTDSGATWSQM 164
Db 292	GGGVORLSVRVEGLANITRVAAGSAHALAVRGDGTVMWCDNEGOL-----GDGSHADR 347
Qy 165	TNFPDVGTIANTPTDTTGYOSDIQGVVWVAFDKSSSSSLGQASKTI----FVGVAADPNPV 220
Db 348	-----FRPVQVFG---LQGITAVSGGRSHSMALQGDGTVRWANGYGYQLGDGT 393
Qy 221	FWSRD-----GGATWQAVPGAP 237
Db 394	LTSRVPALIPGLTGIQALHPSHLHVLAHADGTLRGWGYNRFQQLGLGAAGSAVPPVQV 453
Qy 238	TGFTPHKGVDPVN---HVLITATSN-----GGPYDGSSGDWVKFVSTSGTWT-RISP- 287
Db 454	RGI-----GRYDRLSVGRHAHTLMVRADGTVMWANGENGSGQLGD-----GTSHRTGPV 501
Qy 288	---VPSTDANDYFGYS-----GLTIDRQHPNTI-----MVA 316
Db 502	SVPQVPCVRVSAAGIQHSLALACDGTVMWANGANRGQLDGTTPRVTPLGLEGLRGVVA 561
Qy 317	TQISWPPDTITFRSDCGATWTRIDWTSYPNRSRLRVLDISAPMWTFGVQNPVPVPS 376
Db 562	--LAAGDASVALRADGS-----VWSWGG-----NASGQLGDGTLTDRALPA- 601
Qy 377	KLGMMDAMAIDPFNSDRMLYGTGATLYATNDLTWK-----DSGGQIHAPM-VKG 426
Db 602	HVEHLSNAVPVPVGETHALAVGEDGALWA-----WGANGSGQLGDGTAPSLTPVRVK 655
Qy 427	LEETAVNDLISPPSGAPLIISALGDLGFTHADVTAVPSTIF--TSPVFTTGTSDVYAE 483
Db 656	LERVA-----SVAAGRAFSVAVRDDG-----TAWMGTPSGQLGDGTNHARS 698
Qy 484	NPSIIVRAGSFDPSSQPNDRHVAFTSDGKKNWFG-----SEPGG----- 523
Db 699	VPNQVTSLKGRAMSAAGAHVVALSAD-GTVWTWGDNTLQQLGSGSSSTGMWPRQVPEL 757
Qy 524	-----VTTGG--TVAASADGSRFWAPGDGPQVYVAVGNSWAAASQGVYANAIQRSDR 576
Db 758	RGVEAVAGEQFTVAVLLDGTAAWGSNEYGQ-----LGDGKTGQLTPMAVKEGPK 810
Qy 577	VNP--KTFYALSNGTFYRSDGGV-----TFQPVAAGLPSSGAVGMFHAVPGKEDLW 628
Db 811	IRPVRTVRWQGHAVVQMSDGTVOIWGDNTFGQLGDGTTTTRRAVPLTVYGLASV----- 865
Qy 629	LAASSGLYHS---TNGSGSSSAITGVSSAVNVEFGKSAPGSSYPA-----VFVV 674
Db 866	VAVSSGAWHSIALSLDGTVWA--WGAN-----GFGQLGDGTSTPTFPVQVAGLDVVAI 918
Qy 675	GTGGVTGAYRSDCGTTTWVLINDDQHQYGNWQG 708
Db 919	GSQGYHALAVCSDGSVMTW-----GYNAFGQ 944

RESULT 4

RESULT 4
US-09-902-540-9932
; Sequence 9932. Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.

RESULT 5
US-08-410

```

; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 9932
; LENGTH: 1300
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-9932

```

[illegible]

RESULT 5
US-08-410-784A-5

; Sequence 5, Application US/08410784A

; Patent No. 5912413

; GENERAL INFORMATION:

; APPLICANT: MYERS, ALAN M.

; APPLICANT: JAMES, MARTHA G.

; TITLE OF INVENTION: ISOLATION OF SUL, A STARCH DEBRANCHING

; TITLE OF INVENTION: ENZYME, THE PRODUCT OF THE MAIZE GENE

; TITLE OF INVENTION: SUGARY 1

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Weingarten, Schurgin, Gagnebin and Hayes LLP

; STREET: Ten Post Office Square

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/410,784A

; FILING DATE: 24-MAR-1995

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Heine, Ph.D., Holliday C

; REGISTRATION NUMBER: 34,346

; REFERENCE/DOCKET NUMBER: ISU-002XX

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-542-2290

; TELEFAX: 617-451-0313

; TELEX:

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 772 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE: N-terminal

; ORIGINAL SOURCE:

; US-08-410-784A-5

Query Match

Best Local Similarity 19.8%; Score 162.5; DB 1; Length 772;

Matches 162; Conservative 90; Mismatches 272; Indels 293; Gaps 42;

Qy 51 WIPLLDVGNNWGCYN---GVVS---IAADPINTNKVAAVGMVMTNSWDNDGAILRSSD 104

Db 108 WGP--NNPYASNWKGSGAGCVSDVDANGDRFNENKLLLDPEYAEBSQDP----- 155

Qy 105 QGATWQITPLPFLKGGNMPGRGMGERLAVDPNNNIIYFGAPSGKGLWRSTDSGATWSQM 164

Db 156 -----LNPSNQ-----GNVFASAHYRTDSCI----- 178

Qy 165 TNPDPDGVYIANPDITG-----YQSDIQGVVWVAFDKSSSLG----- 203

Db 179 --YAPKGVLPVPSQSTGTGTRKPAQKDDVIYEVHVRG-----FTEQDTSIPAQVGTYYG 231

Qy 204 ---QASKTIFVGVV-----DPNNPV-----FWSRDCGATWQA 232

Db 232 AGLKASYLASLGVTAVERFLPVQETQNDANDVVPNSDANQYWGVTENYFSPDRRYAANK 291

Qy 233 VPGATG-FIPKGVDPVNHVLIYATSNTPGGYSGDGVWKFVSVTGWTTRISPVST 291

Db 292 AAGGPTAEQAVQAFNAGIKVMDV-----VYNHTABGGTWTSSDPTTAT 338

Qy 292 -----DTANDYF-GVSGLTIDRQHNTI---MVATOISWMPDTIIFRSTD 332

Db 339 IYSWRGLDNATYYELTSGNYFYDNTGICANENTYNTVAQNILVDSVAYWANTM---GVD 395

Qy 333 GGATWTRINDWTSYPNRSLRYVLDISAPWLTFGVQPNPPVPSKLGWMDAEMADPFNS 392

Db 396 GFR-----FDLASVLGNSCLNAVHASA-----PNCPCNGGYNFDDAADSNAI----- 436

Qy 393 DRMLY-----GTGATLY-----ATNDLTWK-DSGGQIHIAPMVKLEETA VNDL 435

Db 437 NRILREFTVRAAGGTWICLNLGPSAATR--TSWVDSRRVVRVWVSVPRLRQANEL 494

Qy 436 ISPPSGAPLISALGDLGFTHADVTAVPSTIFTSVFTTGTSDVYAEINPSIIVR----- 490

Db 495 -----GSMTIYVTQDANDFGSG-----SNLFQSSGRSPWNSINFIDVHDGMTLKD VYSC 543

Qy 491 --AGSFPDSSQPNDRHVAFSTDGKNWFGSGEPGGVTTGCTV---AASADGSRF-VWARG 544

Db 544 NGANNSQASYGSPDG-----GTSTNYSWDQGSAG---TGAAVDQRAARTGMAFEMLSAG 596

Qy 545 DP-----GQPVVYAVGFGN-----SWAASQG---VPANAQI RSDR---VNP K 580

Db 597 TPLMQGDEYLRTOCNNNAYNLDSANWLTYSWTTDQSNFTYFAQLIRSRARHIPRES 556

Qy 581 TFYALNSGTFTYRSTDGVTQFPVAAAGLPSGAGVGMFHAVPGKEGDLMLAASSGLYHSTN 640

Db 657 SWYSGSQTWY-----QP-----SGAV-----ADSNYWNNTS 683

Qy 641 GGSSWSAITGVSSAVNVGFGKSAPOSSYPAPVWV--GTIGGVTGAYRSDDCGTTWVLND 698

Db 684 NYAIYAIANGPSL-----GDSNDSIYVAYNGWSSSVTFTLPAPPSGTQWYRVTD 732

Qy 699 DQHOYGNWQAITGDHANLRVYIGTNGRGIYVYGDIG 735

Db 733 T-----CDWNDGASTFVAPGSETLIG--GAGTTYGCGG 763

RESULT 6

US-09-248-796A-14123

; Sequence 14123, Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 14123

; LENGTH: 1060

; TYPE: PRT

; ORGANISM: Candida albicans

US-09-248-796A-14123

Query Match

Best Local Similarity 4.0%; Score 159.5; DB 2; Length 1060;

Matches 147; Conservative 68; Mismatches 267; Indels 223; Gaps 33;

Qy 7 TWSNVAIGGGGFDGIVFNEGAPGILYVYRTDIGG-----MYRWDAAANGRW--IPL-LDWV 58

Db 142 TYKNVPAGYRPFVDAYI-----SATDWSYTLVSANEYTCAGGVQWRAPPTLRWT 191

Qy 59 GWNH--HGYNQGVSIADDPINTNKVAAVGMVMTNSWDNDGAILRSSQOGATWQI--TPLP 115

Db 192 GYRNSAGSNGSIVVAITRTVTDSTAVT---TLFPDEN-----RDKTKTIELKPTP 241

Qy 116 FK-----LGGNMPGRGMGERLAVDPNNNIIYFGAPSGKGLWRST-DSGATW 161

Db 242 TTTTTSYGVTSYLTJKTAP---IGETATVIV---DIPYHTTTTTSKWTGTTTSTTH 295
Qy 162 SQTNPFP---VGYIANTP-DTGYQSDIOGVVWAFDKSSSLGQASKTIFVGVADPN 217
Db 236 TNPSTDIDTIVQVLEPNPTVTTEYSQ-----SPATTTITGPPGNTDVLIREPP 348
Qy 218 NPV-----FNSRD---GGATWQAVCAPTGFPHKGVFDPVNHVL-----YIATSN 261
Db 349 NHTVTTEYSYVTTSTFAPPGTDSVI-----IKEPPNPTVTTEYWSYTTT 404
Qy 262 GGPYDG-----SSGDVWKFVSSTGWTTRISPVSTD-----TAN 295
Db 405 TAPPGTDTVLIIRPPNHTVTTEYSQSYTT-TTIVIAPEGGTDVLIIRPPNPTVT 463
Qy 296 DYGVSGLT-----IDRQHPNTIMVATQISWV-----PDT 325
Db 464 EYWSQYATTTTITAPPGTDTVLIIRPPNHTVTTE--YWSQSYATTTTITAPPGTDT 521
Qy 326 IIFSTGGATWRIWDTSYPNRSLRYLDISAEPLTFGV-----QNPVPSPKL--- 378
Db 522 VLIIRPPNHTVT---TEYWSQSYTTTIVIAPEGGTDVLIIRPPNPTVTTEYSQ 576
Qy 379 -----GMDDEAMATDPNSDRMLYGTGATLYATNDLTKWDSGGQIHIA----- 421
Db 577 SYATTTITAPPGTDTVLIIRPPNHTVTTEYSQSYATTTTITAPAGTDSVLIIRPP 636
Qy 422 -PMVKGLE-----ETAVNDLISPPSAPGLISALGDLGGFTHADVTAPSTIFTSVFTG 475
Db 637 NPTVTTEYSQSYTTTNTVTAPEG-----TDSVLIIRPPNPTV 676
Qy 476 TSDVYAEINPSIIVRAGSFDPSQPNDRHVAFTDGGKNVFGSGPGVT----- 525
Db 677 TTTEYSQSYATTI-----TVTATPGGTDVLIIRPPNPTVTTEYSQSY 721
Qy 526 --TGTVAASADGRFVWAPGDPQPVVYAVFGNSWAASQVPA 568
Db 722 YATTTTAPPGTDTVLIIRPPNPTVTTEYSQSYATTTTVA 766
RESULT 7
US-09-346-237-8
; Sequence 8, Application US/09346237A
; Patent No. 6265197
; GENERAL INFORMATION:
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Svendsen, Allan
; TITLE OF INVENTION: Starch Debranching Enzymes
; FILE REFERENCE: 5629.200-US
; CURRENT APPLICATION NUMBER: US/09/346,237A
; EARLIER FILING DATE: 1999-07-01
; EARLIER FILING DATE: 1998-07-02
; EARLIER FILING DATE: 1998-07-02
; EARLIER FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Favobacterium odoratum
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(774)
; OTHER INFORMATION: Isoamylase
US-09-346-237-8
Query Match 3.9%; Score 158.5; DB 2; Length 774;
Best Local Similarity 18.8%; Pred. No. 0.0035;
Matches 158; Conservative 103; Mismatches 265; Indels 313; Gaps 44;
Qy 76 INTNKVWAA-----VGMVTNSWDPNDGAILRSSDQATWQITPL 114
Db 35 INPNKLGAAVDATKANVTFKYSSKATRIEILYSTATGSAEKAKYVMTNSGGIWSVT-I 93

Qy 115 PFKLGNMPCRMGERLAVDPNNDNLLYFGAPSGKGLWRSTDSGATWSQMTNPPDVGTI 174
Db 94 P---TSTLSGQGLG-----GTLYYGYRAWGNW---PYNASWTGSSSL----- 130
Qy 175 ANPTDITGYQSDIQG-----VWVAFDKSSSLGQASKTIFVGVADPNNP 219
Db 131 -----GFISDVDAAGNRFNPNKLLSDPYALELSHDPITATM-----TNGS 170
Qy 220 VFWSDGGATWQAV---PGAPTGF-----PHKGVFDPV---NHLVIATSNIG 262
Db 171 IYAS---GATYNIOSGSAKGIVLADGTQATGTFKTRALKDDVVEAHVRLGTWMDTS 227
Qy 263 -----GPYDSSGDVWKFVSSTGWTTRISPV-----PSTDANDYFGYSLTI 305
Db 228 ITAAVRYGTGAGLKAALAAALGVTAEFLPVQETQNDTNDNDPSSSTSGDNYWGYMTLNY 287
Qy 306 ---DRQ-----HPNTIMVATQISWVPTTIIIFRSDGGATWTRIW 341
Db 288 FAPDRRYAYDKTPGGTREFKEMVKAFHDNGIKVLVD-----VVNHTGEGAMSPT- 339
Qy 342 DMTSPNRSRLR-----YVLDISAE-PWLTGCVQNPVPVS-----PKLGMDEAM 385
Db 340 DKTTNITSFRGLDNPYYLSLTADFQNSWDNTGVGNYNTRNTIAONLIVDSLAWRDKL 399
Qy 386 AIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPMVKLEETAVNDLIS--PP--- 439
Db 400 GVDGYREFDLASVLGNSCQHCFCNFKWDAG-----NALNRIVAEI,PPRPAT 445
Qy 440 ---SGAPLIS---ALG-----DLGGF-----THADVTAVPSTIFTSVFTTGTSDVYA 481
Db 446 GSGVDLIAEPAWAGSYQVGGFSGWAEWNGAYRDVVRQAQNLKLSVAITTG----- 499
Qy 482 ELNPSIIVR-ACSFPPSSQPNDRHVAFTDGGKNVFGQ---SEPGVTTTGTVAASADGS 537
Db 500 ---QMATRFAGSSD-----LYGDDGRKPHSVNFIATHDGFTLKDLYSCSKNN 545
Qy 538 RFVW--APGDPQPVVYAVFGNSW-----AASQ-----GVP-----ANAQ 571
Db 546 NOVWPGFSDGGE-----DNNNSWDQGGIADQKAAANGMALMMLSAGVPMIVGDEA 599
Qy 572 IRS-----DRVNPRTFY-----ALSNGTFFRSTD-G 596
Db 600 LRSMCNPNPNLDSSANLWNSRTTQNNFQSFASKAMIAFRKAHPALRPANFYSSVDNN 659
Qy 597 GVTFOVAAGLPSSGAVGMF-----HAPVQK-EGDLMLAASSGLYHSTNGSSWSAIT 649
Db 660 GNVMEQLRWFKPDGGVADATYFENDANNHAIARIDGSEFGDTASAIYVAHN---AWSA-- 714
Qy 650 GVSSAVNVGFGKSAPGSSYPVAVVVTGICGVTGAYSDDCGTTWVLINDDOHOYGNWQ 708
Db 715 -----QVNFTLPWPGAGKSWYRVTDTGWAEGASQVQAPGSE-ALVGGENTAYGLGCR 766
RESULT 8
US-09-489-039A-9813
; Sequence 9813, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9813
; LENGTH: 688
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9813

```

Query Match      3.8%; Score 157.5; DB 2; Length 688;
Best Local Similarity 22.6%; Pred. No. 0.0035;
Matches 111; Conservative 53; Mismatches 184; Indels 143; Gaps 24;

QY 60 WNNWG--YNGVSTIADPINTNKVAAVGMVYNSWDNDGAILRSSDQATQITPLPFK 117
Db 171 WAWGNTHDRFAALDQINKN-----VNLQVAVAHATGDIPOSNGAEDQNTPLQ-- 224
QY 118 LGGNMPGRGMBERLAVDPNNNDNLYFGAPSGKGLWR--STDGATW-----SOM 164
Db 225 -----IGDTLYCTPYSKVLALDSDGKWKRYDSKSSPNWQRCRGLGYADSQA 275
QY 165 TNPFDVGT-----YIANPTDT-----TGVSQDI-----QGVVWVAFD 196
Db 276 QTAPASGTQPAACSRRLFLPTIDARLIAIDATGKLCENFGDGGIVDLVSGMGEVKAGYY 335
QY 197 KSSSLGQASKTIFVG--VAD--PNNPVFWSRDGGATWQAVCAPTFPHKG-----VF 247
Db 336 QQTSTPLVAGNVVVGVRADNYSVTGEP-----PGVVRADFVHTGKLAWAW 381
QY 248 DPNVHLYIATSNTPGYPDSSGDVWKFVSVTSGT---WTRIS-----PVSTDTAND 296
Db 382 DPGNPAL-----TGVPPEQ-----TYTRGTPNVWSAMSVDKLNLIYLPFGNATPD 428
QY 297 YFGVSGLTIDRQHPNTIMV--AT--QISWMPDTIIFRSTDCGATWTWTRIDWTSYPNRSR 352
Db 429 FFGERTALDDKYSSSIVAVDAITGQVRHFQT-----THDLWDF-DLPSQPLL 477
QY 353 Y-----VLDISAEPWLIFGVQNPFPVSPKLGWDEMAIDPFN-----SDRM 395
Db 478 YDLPDGGKGTTPVLVQTSKQGMIFMLNRETGEPAKV-----EERPVPAGNVKGERYSPTQ 533
QY 396 LYTGATLYATNDLTWDSGCGQIHIAPMVKGLE--ETAVNDLISPPSGAPLISALGLGG 453
Db 534 PYSVGMPMIGNQTLTSDMWGATPIDLLLCRIQFKEMRHOGVFTPPGEDRSLOFPGLSG 593
QY 454 FTHADTVAPVS 464
Db 594 MNWGSVSLDEN 604

RESULT 9
US-09-252-991A-22301
; Sequence 22301, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22301
; LENGTH: 1548
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22301

Query Match      3.8%; Score 154; DB 2; Length 1548;
Best Local Similarity 20.7%; Pred. No. 0.019;
Matches 188; Conservative 87; Mismatches 270; Indels 364; Gaps 53;

QY 73 ADPINTNKVAAVGMV---TNSWD-----PNDGAILR----- 101
Db 319 SDGVROGRNWALAEMLPAGSQWSMRLVAGADLAADNRLVRPDSASLADLADTHYQAKI 378
QY 102 -----SSDQATQITPLPFKLGNNMPGRGMRGLAVDPNNDNLYFGAPSGKGLWR 153

```

RESULT 10

```

US-09-136-574A-44
; Sequence 44, Application US/09136574A
; Patent No. 6294366
; GENERAL INFORMATION:
; APPLICANT: Farrington, Graham K.
; Anderson, Paige
; Gibbs, Moreland
; Berquist, Peter
; Daniels, Roy
; Morgan, Hugh W.
; Williams, Diane P.
; TITLE OF INVENTION: Compositions and Methods for

```

```

Db 379 EQSSGLVFTDQADWGLTP-----GTPVDESNEWI-----CGLGPPYC 416
QY 154 STDSGATWSOMTNPFPDVGTYIANPTDITTYQSDIQGVVWVAFDKS--SSSLGQASKTIFV 211
Db 417 AEPRTWAP-----GNLYGMPAGTAIGBDLW--WCSVDPLSLCILENGLKT----- 460
QY 212 GVADPNPNPVF-----MSRDG--GATWQAVPGAPTFPHKGVFDP 249
Db 461 -VVTQNLFSVLRTGTGDLDLASAGNLTQWSPYGVYTAGTQAADVA-TGFNQPRGLEF- 517
QY 250 VNHVL-----YIATS-----NTGGPYD--GSS--GDVWKFVSYSGTWTTRISVP 289
Db 518 -GSVLGAGGADYEVLSTSQYQAWPEHGGNLDIAVGGDVGVGDQWAEKLTSS--DIRPLP 574
QY 290 STDANDYFYGSLTIDRQHPNTIMVATQISWMPDTIIF-RSTDCGATWTRIDWNT----- 344
Db 575 PSAAVGNWLMROG-SADREGVPT-----AWWNFGSVYRGAEGDAPY--LVGTGTGFT 624
QY 345 -----SYPN-----PSLYVLDISABEPLWT----- 364
Db 625 LGGNLSMRTGGDAGNIAPRGDGSIPSSGNLNPSSQGLVAVAGTGRITSDCALQLGGG 684
QY 365 -----FGVQPNPVPSPKLGWDEMAIDPFNSDRM--LYGTG-----AT 402
Db 685 DLNVRIGGEVNP-----SREARATQYSSSGFDGLYSGGTHDLOGALINLRGSAS 735
QY 403 LYATNDLTWDSGCGQIHIAP---MVKGLETAVNDLISP-----PSGAPLIS----- 446
Db 736 LY-----SGALGIDPRYDILLRDPAEVRSRDAFSPTLASSTGGTLVAVGDTGMRL 786
QY 447 -ALGD--LGFTHADVTAVPSTI--FTSP-----VFTGTSTVDYAEINPSII 488
Db 787 ETRGDLVLGVTDPRGVGNPTVGTAPDGSVYQGGIGWFSLWTAHTSID-----LF 839
QY 489 VRAGSFDPSSQ-----PNDRHVAFSTD-----GKKNWFQSGE 520
Db 840 AAGCNLTPTQLVEATNAIPMAGRNLSPSDGRFTYPSIVRAAAPRGSIYLPSSGVMGV 899
QY 521 PGVYTT---GGTVAASADGSRFFWAPGDPGPVVYVAVCGFNSWAASQGVANAQI----- 572
Db 900 SLNVSTTYSLLAPLSINGELELLA-GDS-----IYAGGY--SVQSRGADPANLPSIWTPA 952
QY 573 ---RSDR--VNPKTFYALNSGTFRSTGVTQFQVAAGLPSSGAVGMFHAVPGKEGDL 627
Db 953 FAGYSDAALNP-----IAGNG-----SPDGN--PAVIGGLP-----LFYFGPDSAASL 994
QY 628 WLAASSGLYHSTNGSSWSAITGVSSAVNVGVFGKSAPGSSYPVFFVGTIGVTVGAYRSD 687
Db 995 ARDLQPARFYALTGD-----IVGLNSGAQIRFGQA-----GNRAGQTWYE 1035
QY 688 DCGTTWVLINDDQHOYGN-WQOAI-----TGD-----HANLRVYIGTNGRGIVY 731
Db 1036 GAGPVWNRAGRDIIVASGTPLGQRISAFPSQISTDASFTGNLFVHDDPNDLSLVQAGRDILY 1095
QY 732 GDIGGAPSG 740
Db 1096 GNFVAGPG 1104

```

[illegible]

```

QY 196 DKSSS-----SLQASKTIFGVADPNPNPFW-----RDGATWQAVPG- 235
Db 1099 GSGSVHRYSLPSYAVGAKYEETAMIG-CSPGLEFSSDGLRFPVDPAGSETAAVYGL 1157
QY 236 -APTGFPHKGVDPV-----NHVLIATSTNGG----- 264
Db 1158 AAPYGI-----GEAEPLPLFLGVGAEBEATLSPDGRHILV-----PGRGLSQYSLFSTNL 1208
QY 265 -----YDSSGD--VWKFVSVTGWTWTRISPVSTDTAN-----DYFGYSG----- 302
Db 1209 ELCAEPRGIDGSCBEDIYAPE-SPGRGEGVSLAASITAADCPGIGELHGFAGPMPAPV 1267
QY 303 ---LTIDRQ-----PNTIWTAT---QISWWDPT----- 1139
Db 1268 MEQVTLDSREGTLRVLDRIVDVTVPYKMWVEDSDGQSTTLANSTLLNAENSNILFR 1327
QY 330 STDGGA-----TWTRIW---DWTGYPNR-----SLRYVLDISAEPWLTFCVQ 368
Db 1328 LDDAAAGKISGTSFVFRWSSPFLGTGATRPHLGGVRLADIYDASG----- 1379
QY 369 PNPVPSPK-LGMWDEAMADPPNSDRMLYGTGA-----TLATNDLTWDSGGQIHIA 421
Db 1380 ---VSPSGIEFSDGGRM-----FVTGIGTPGINIFTLSAPFDITLPHKSGSTNIG 1428
QY 422 PMVGLBETAVNDLISPPSGAPL-----ISALGD-----LGGFTH-----ADVT-A 461
Db 1429 ---GL---SVSDLAFAFANGNSLTVLVDVGVRLVYALGDDVNVVTGTTQKFRITLDTTQ 1481
QY 462 VPSTFTSP-----VFTCTSDVYA---ELNPSIIVRAGSDPSS-----QP 500
Db 1482 IPNSIYTPDGLSFVAYDRIIDYVLGSPNDISSTETIIPYSLPRP-----DPPTGMDFT 1538
QY 501 NDRHVAFSTGCKWFOGSPGGVTTGG-----TVAASADGSRFV-----WAPGDPG-- 547
Db 1539 DGRMFLSTENGIDQYLLSEPAFTTSVFLRTIDGGAEGIRFVNGRGLFVPGADGII 1598
QY 548 --QPVTVAVGNSW--AASQGVAPANAQIRSDRVNPKTFYALSNGTFYRSTDDGGVTFQPV 603
Db 1599 QRHELIYPYGASTSILETVRDGV-----TDGPGFGENPA 1631
QY 604 AAGLPSSGAVGMHFAVPGKECDLWLAASSGLYHSTNGGSSWSAITGVSSANVNVGFGKSA 663
Db 1632 AGEIRLAGT-----FNASDN-----VQSPSGIEFSGDGTGMF--VTGFGAAGVNEFSLSA 1679
QY 664 P-GSSYPAVFVVGTTGG---VTGAYRSDCCGTTWVLNDD 699
Db 1680 PFTTLFVHVELHDIGQPAVDLAFAD--GRTLLLLAAD 1717

```

RESULT 12

```

US-09-346-237-4
; Sequence 4, Application US/09346237A
; Patent No. 6265197
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Svendsen, Allan
; TITLE OF INVENTION: Starch Debranching Enzymes
; FILE REFERENCE: 5629.200-US
; CURRENT APPLICATION NUMBER: US/09/346,237A
; CURRENT FILING DATE: 1999-07-01
; EARLIER APPLICATION NUMBER: PA 1998 00868
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: 60/094,353
; EARLIER FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 776
; TYPE: PRT
; ORGANISM: Pseudomonas amylocleromosa
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(776)

```

; OTHER INFORMATION: Isoamylase
US-09-346-237-4

Query Match
Best Local Similarity 3.8%; Score 151.5; DB 2; Length 776;
Matches 171; Conservative 89; Mismatches 274; Indels 291; Gaps 47;

```

QY 27 GAPGILYVRTDIGMYRWDAANGRIPLLDVGMNMGVGVVSIADPINTNKVWAAV 86
Db 96 GITGAVY-----YGYRWGPNWYASNGKGSQAGFVSDVANGDRFNENKLLDPY 147
QY 87 MYTWSWPDNCAILRSSDQATWITPLPFLGGMNPGRCGERLAVDPNDNLLIFGAP 146
Db 148 AOEVSQDP-----LNPNSQ-----NGNVFASGASR-----TTDSGIY--AP 182
QY 147 SGKGLWRSTDSGATWSQMTNFPDV--GTYIANPTD--TTGYOSDIOGVVWVAFDKSS--SS 201
Db 183 KGVVLVSTQSTGTKPTRAQKDDVIYEVHVRGTFEODTISIPAYRGTYTGAGLKASYLAS 242
QY 202 LQASKTIFVGVAD-----PN---NPVFW-----SRDGGATWQAVPGAPTG-F 240
Db 243 LG-VTAVEFLPVQETQNDANDVVPNSDANQYVGMTEYFSPDRRYAYNKAAGGPTAEF 301
QY 241 IPHKGVPDPVNHVLIATSTNGGPGYDGSDDVWKFVSVTGWTWTRISPVST----- 291
Db 302 QAMVQAFHNAGIKVYMDV-----VYNHTAEGGTWTSDDPTATTIYSWRGLDN 348
QY 292 -----DTANDYF--GYSGLTIDRQHPNTI---MVATOISWMPDITIIFRSTDCGATWTRIW 341
Db 349 ATYYELTSGNQYFVDNTGIGANFNTYVAONLIVDSLAYWANTW---GVDGFR-----F 400
QY 342 DWTS-YPNRSRYVLIDISAEPLWTFGVQPNPVPSPKLGWMDMAID-----PFNSDRML 396
Db 401 DLASVLGSLNGAYTASA-----PNCNGGYNFADAADSNVAINRILREFTVRPA 451
QY 397 YGTGATLYATNDLTWDSGGQIH-IAPMVKGLEE-----TAVNDLISPPSGAPLI 445
Db 452 GSGGLDLFA---BPWATGNSYQLGGPPQGWSEWNGLFRDSLRAQONEL-----GSMTI 502
QY 446 SALGDLGGFTHADVTAVPSTIFTSPVFTTGTSDVYAEINPSIIVR-----AGS 493
Db 503 YVIQDANDFSGS-----SNLFQSGSRPMSINIFDVHDMGTLDKDVYSCGNANNSQWP 556
QY 494 FDPSSQPNDRHVAFTDGGK---NWFGSPGGVTTGGTV---AASADGSRF-VWAPGD 545
Db 557 YGPS-----DGTSTNYSWDQMSAG---TGAAVDORRAARTGMAFEMLSAGT 601
QY 546 P---GQPVVYAVGFGN-----SWAASQGVAPANAQIRSDR-----V 577
Db 602 PLMQGGDEYLRQLCANNAYNLDSANWLTYSWTTDQ---SNFYTFAQRLLIAFRKAHPAL 658
QY 578 NPKTFYALNSGTFYRSTDDGVTFFQPVAAAGLPSGAVG-----VMHFAVPGKEGD 626
Db 659 RPSSWYSGSLTWY-----QP-----SCAVADSNYWNNTSNYAIAYAINGPS--- 700
QY 627 LWLAASSGLYHSTNG-----GSSWSAIT-----GVSSAVNVGFGKSA 665
Db 701 --LGDSNIYVAYNGWSSSVFTTLPAPSGTQWYRVVTDTCMDNDGASTFV-----APG 751
QY 666 SSYPAVFVVGTTGGVGTGAYRSDCCGTTWVLNDDQHOYGNMGOAI 710
Db 752 SE-----TLIGG-----AGTT-----YGCQGSQ 770

```

RESULT 13

```

5457037-3
; Patent No. 5457037
; APPLICANT: TOGNONI, ANGELO;CARRERA, PAOLO;CAMERINI, BARBARA;
; GALLI, GIULIANO;LUCCHESI, GIUSEPPE;GRANDI, GUIDO;DI GENNARO, CARLO
; TITLE OF INVENTION: CLONING OF THE GENE CODING THE ISOAMYLASE
; ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME
; NUMBER OF SEQUENCES: 7
; CURRENT APPLICATION DATA:

```


5457037-5	APPLICATION NUMBER: US/08/1,797	3.7%; Score 148.5; DB 6; Length 751;	Query Match	27	GAPGILYVRTDIGMYRWDAAANGRWIPLLDVGMNNGVYVSIADPINTKNKWAAVG	86
	FILING DATE: 08-JAN-1993	20.7%; Pred. No. 0.018;	Best Local Similarity	70	GITGAVY-----YGRANGPNWPYASNMKGKQAGFVSDVDANGDRFNPKNLLDPY	121
	PRIOR APPLICATION DATA:	Matches 171; Conservative 88; Mismatches 275; Indels 291; Gaps 47				
	FILING DATE: 19-AUG-1991					
	APPLICATION NUMBER: 224,114					
	FILING DATE: 25-JUL-1988					
	SEQ ID NO:5					
	LENGTH: 751					
5457037-5						
	APPLICANT: TOGNONI, ANGELO;CARRERA, PAOLO;CAMERINI, BARBARA;					
	GALLI, GIULIANO;LUCCHESI, GIUSEPPE;GRANDI, GUIDO;DI GENNARO, CARLO					
	TITLE OF INVENTION: CLONING OF THE GENE CODING THE ISOAMYLA					
	ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME					
	NUMBER OF SEQUENCES: 7					
	CURRENT APPLICATION DATA:					
	APPLICATION NUMBER: US/08/1,797					
	FILING DATE: 08-JAN-1993					
	PRIOR APPLICATION DATA:					
	APPLICATION NUMBER: 749,621					
	FILING DATE: 19-AUG-1991					
	APPLICATION NUMBER: 224,114					
	FILING DATE: 25-JUL-1988					
	SEQ ID NO:5					
	LENGTH: 751					
5457037-5						
	APPLICANT: TOGNONI, ANGELO;CARRERA, PAOLO;CAMERINI, BARBARA;					
	GALLI, GIULIANO;LUCCHESI, GIUSEPPE;GRANDI, GUIDO;DI GENNARO, CARLO					
	TITLE OF INVENTION: CLONING OF THE GENE CODING THE ISOAMYLA					
	ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME					
	NUMBER OF SEQUENCES: 7					
	CURRENT APPLICATION DATA:					
	APPLICATION NUMBER: US/08/1,797					
	FILING DATE: 08-JAN-1993					
	PRIOR APPLICATION DATA:					
	APPLICATION NUMBER: 749,621					
	FILING DATE: 19-AUG-1991					
	APPLICATION NUMBER: 224,114					
	FILING DATE: 25-JUL-1988					
	SEQ ID NO:5					
	LENGTH: 751					
5457037-5						
	APPLICANT: TOGNONI, ANGELO;CARRERA, PAOLO;CAMERINI, BARBARA;					
	GALLI, GIULIANO;LUCCHESI, GIUSEPPE;GRANDI, GUIDO;DI GENNARO, CARLO					
	TITLE OF INVENTION: CLONING OF THE GENE CODING THE ISOAMYLA					
	ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME					
	NUMBER OF SEQUENCES: 7					
	CURRENT APPLICATION DATA:					
	APPLICATION NUMBER: US/08/1,797					
	FILING DATE: 08-JAN-1993					
	PRIOR APPLICATION DATA:					
	APPLICATION NUMBER: 749,621					
	FILING DATE: 19-AUG-1991					
	APPLICATION NUMBER: 224,114					
	FILING DATE: 25-JUL-1988					
	SEQ ID NO:5					
	LENGTH: 751					
5457037-5						
	APPLICANT: TOGNONI, ANGELO;CARRERA, PAOLO;CAMERINI, BARBARA;					
	GALLI, GIULIANO;LUCCHESI, GIUSEPPE;GRANDI, GUIDO;DI GENNARO, CARLO					
	TITLE OF INVENTION: CLONING OF THE GENE CODING THE ISOAMYLA					
	ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME					
	NUMBER OF SEQUENCES: 7					
	CURRENT APPLICATION DATA:					
	APPLICATION NUMBER: US/08/1,797					
	FILING DATE: 08-JAN-1993					
	PRIOR APPLICATION DATA:					
	APPLICATION NUMBER: 749,621					
	FILING DATE: 19-AUG-1991					
	APPLICATION NUMBER: 224,114					
	FILING DATE: 25-JUL-1988					
	SEQ ID NO:5					
	LENGTH: 751					
5457037-5						
	APPLICANT: TOGNONI, ANGELO;CARRERA, PAOLO;CAMERINI, BARBARA;					
	GALLI, GIULIANO;LUCCHESI, GIUSEPPE;GRANDI, GUIDO;DI GENNARO, CARLO					
	TITLE OF INVENTION: CLONING OF THE GENE CODING THE ISOAMYLA					
	ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME					
	NUMBER OF SEQUENCES: 7					
	CURRENT APPLICATION DATA:					
	APPLICATION NUMBER: US/08/1,797					
	FILING DATE: 08-JAN-1993					
	PRIOR APPLICATION DATA:					
	APPLICATION NUMBER: 749,621					

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 14:20:38 ; Search time 142.551 Seconds
(without alignments)
2169.009 Million cell updates/sec

Title: US-09-917-376-3

Perfect score: 4036

Sequence: 1 ATTQPYTWSNVAIGGGFVD.....YIGTNGRGIIVGDIIGAPSG 740

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.Main:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pap.*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pap.*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pap.*
- 4: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pap.*
- 5: /cgn2_6/ptodata/1/pubpaa/US108_PUBCOMB.pap.*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4036	100.0	740	US-09-917-376-3	Sequence 3, Appli
2	4036	100.0	740	US-09-917-376-6	Sequence 6, Appli
3	4036	100.0	740	US-10-155-400-3	Sequence 3, Appli
4	4036	100.0	740	US-10-155-400-6	Sequence 6, Appli
5	4036	100.0	957	US-09-917-376-1	Sequence 1, Appli
6	4036	100.0	957	US-10-155-400-1	Sequence 1, Appli
7	2429.5	60.2	882	US-10-156-761-9395	Sequence 9395, Ap
8	1680	41.6	726	US-09-917-376-7	Sequence 7, Appli
9	1680	41.6	726	US-10-155-400-7	Sequence 7, Appli
10	1579	39.1	838	US-10-420-191-2	Sequence 2, Appli
11	1573.5	39.0	818	US-10-026-994-2	Sequence 2, Appli
12	1405	34.8	739	US-10-156-761-10111	Sequence 10111, A
13	1117	27.7	789	US-10-395-241-14	Sequence 14, Appl
14	1117	27.7	812	US-10-395-241-12	Sequence 12, Appl
15	1114	27.6	826	US-10-395-241-18	Sequence 18, Appl
16	841.5	20.8	555	US-09-927-827-47	Sequence 47, Appl
17	238	5.9	2468	US-10-246-330-4	Sequence 4, Appli
18	238	5.9	2468	US-10-282-122A-66335	Sequence 66335, A
19	199.5	4.9	2435	US-10-282-122A-47453	Sequence 47453, A
20	193	4.8	1465	US-10-369-493-13955	Sequence 13955, A
21	181	4.5	1308	US-10-282-122A-44566	Sequence 44566, A
22	172.5	4.3	783	US-10-369-493-14105	Sequence 14105, A
23	170	4.2	1074	US-10-282-122A-50616	Sequence 50616, A
24	168.5	4.2	1439	US-10-282-122A-44567	Sequence 44567, A
25	167.5	4.2	1119	US-10-245-802-12	Sequence 12, Appl
26	165.5	4.1	613	US-10-369-493-11181	Sequence 11181, A
27	163.5	4.1	1289	US-10-282-122A-44901	Sequence 44901, A

ALIGNMENTS

RESULT 1

US-09-917-376-3
; Sequence 3, Application US/09917376
; Publication No. US20040038334A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOLYTICUS
; FILE REFERENCE: 40197.4US01
; CURRENT APPLICATION NUMBER: US/09/917,376
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
; FEATURE:
; OTHER INFORMATION: Catalytic domain GH74
US-09-917-376-3

Query Match 100.0%; Score 4036; DB 3; Length 740;
Best Local Similarity 100.0%; Pred. No. 4.7e-284;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATTQPYTWSNVAIGGGFVDGIVFNEGAPGILYVRDIDGMYRWDANGRWIPLLDVWVGW	60
Db	1	ATTQPYTWSNVAIGGGFVDGIVFNEGAPGILYVRDIDGMYRWDANGRWIPLLDVWVGW	60
Qy	61	NNWGVGVYSIADPINTNKVAAVGMVYNSWDPNDGAILRSSDQGATWQITPLPEKLG	120
Db	61	NNWGVGVYSIADPINTNKVAAVGMVYNSWDPNDGAILRSSDQGATWQITPLPEKLG	120
Qy	121	NNPGRGMGERLAVDPNNDNLIYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDT	180
Db	121	NNPGRGMGERLAVDPNNDNLIYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDT	180
Qy	181	TYQSDIQGVVWVAFDKSSSLQASKTIFVGVADPNNPVFVSRDGGATWQAVPGAPTGF	240
Db	181	TYQSDIQGVVWVAFDKSSSLQASKTIFVGVADPNNPVFVSRDGGATWQAVPGAPTGF	240
Qy	241	IPHKGVDPNVHVLVYATSNTCGPDYDGGSDGVKFSVTSGTWTRISPVGSTDTANDYFGY	300
Db	241	IPHKGVDPNVHVLVYATSNTCGPDYDGGSDGVKFSVTSGTWTRISPVGSTDTANDYFGY	300
Qy	301	SGLTIDRQHPNTIMVATQISWNPDTTIFRSTDCGATWTRIDWTSYPNRSIRLVYLDISAE	360

Db 301 SGLTIDRQHPNTIMVATQISWMPDIIIFRSTDDGATWTRIDWTSYPNRSRLRYVLDISAE 360
 Qy 361 PMLTFGVQPNPPVPSKLGWDEAMAIIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI 420
 Db 361 PMLTFGVQPNPPVPSKLGWDEAMAIIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI 420
 Qy 421 APWKGLLEETAVNDLISPPSGAPLISALGDLGGTHADVTAVPSTIIFTSVPVFTTGTSDVY 480
 Db 421 APWKGLLEETAVNDLISPPSGAPLISALGDLGGTHADVTAVPSTIIFTSVPVFTTGTSDVY 480
 Qy 481 AELNPSIIVRAGSPDPSSQPNDRHVAIFSTDDGKNWFGQSEPGVTTGGTVAASADGSRFV 540
 Db 481 AELNPSIIVRAGSPDPSSQPNDRHVAIFSTDDGKNWFGQSEPGVTTGGTVAASADGSRFV 540
 Qy 541 WAPGDPQPVVYAVFGNSWAASQGVPANAIQIRSDRVNPKTFYALSNGTFYRSTDDGVTF 600
 Db 541 WAPGDPQPVVYAVFGNSWAASQGVPANAIQIRSDRVNPKTFYALSNGTFYRSTDDGVTF 600
 Qy 601 QPVAAGLPSSGAVGMFHAVPGKEGDLWLAASSGLYHSTNGSSWSAITGVSSAVNVGFG 660
 Db 601 QPVAAGLPSSGAVGMFHAVPGKEGDLWLAASSGLYHSTNGSSWSAITGVSSAVNVGFG 660
 Qy 661 KSPAGSSYPVAVVGTGGTGYRSDDCGTTWVLIINDDOHQYGNWGOAITGDHANLRV 720
 Db 661 KSPAGSSYPVAVVGTGGTGYRSDDCGTTWVLIINDDOHQYGNWGOAITGDHANLRV 720
 Qy 721 YIGTNGRGIVYGDIGGAPSG 740
 Db 721 YIGTNGRGIVYGDIGGAPSG 740

RESULT 2

US-09-917-376-6
 ; Sequence 6, Application US/09917376
 ; Publication No. US20040038334A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DING, SHI-YOU
 ; APPLICANT: ADNEY, WILLIAM S.
 ; APPLICANT: VINZANT, TODD B.
 ; APPLICANT: HIMMEL, MICHAEL E.
 ; TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
 ; TITLE OF INVENTION: CELLULOLYTICUS
 ; FILE REFERENCE: 40197.4US01
 ; CURRENT APPLICATION NUMBER: US/09/917,376
 ; CURRENT FILING DATE: 2001-07-28
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 740
 ; TYPE: PRT
 ; ORGANISM: Acidothermus cellulolyticus
 US-09-917-376-6

Query Match 100.0%; Score 4036; DB 3; Length 740;
 Best Local Similarity 100.0%; Pred. No. 4.7e-284;
 Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ATTPYTWNSNVAIGGGFVDGIVFNEGAPGILYVRTDIGMYRWDAAANGRWIPLLDVWG 60
 Db 1 ATTPYTWNSNVAIGGGFVDGIVFNEGAPGILYVRTDIGMYRWDAAANGRWIPLLDVWG 60
 Qy 61 NNWYNGVWSIAADPINTNKVAAVGMVTSNDPNDGAILRSSDQATWQITPLPFLKGG 120
 Db 61 NNWYNGVWSIAADPINTNKVAAVGMVTSNDPNDGAILRSSDQATWQITPLPFLKGG 120
 Qy 121 NMPGRGMRGLAVDPNNNDLILYFCAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDT 180
 Db 121 NMPGRGMRGLAVDPNNNDLILYFCAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDT 180
 Qy 181 TGQSDIQGVVWVAFDKSSSLGQASKTIFVGVADPNPFWSRDGGATWQVPGATPF 240
 Db 181 TGQSDIQGVVWVAFDKSSSLGQASKTIFVGVADPNPFWSRDGGATWQVPGATPF 240

Qy 241 IPHKGVPDPVNHVLIATISNTGCPYDSSGDVWKFSTSGTWTTRISPVPSTDTANDYFGY 300
 Db 241 IPHKGVPDPVNHVLIATISNTGCPYDSSGDVWKFSTSGTWTTRISPVPSTDTANDYFGY 300
 Qy 301 SGLTIDRQHPNTIMVATQISWMPDIIIFRSTDDGATWTRIDWTSYPNRSRLRYVLDISAE 360
 Db 301 SGLTIDRQHPNTIMVATQISWMPDIIIFRSTDDGATWTRIDWTSYPNRSRLRYVLDISAE 360
 Qy 361 PMLTFGVQPNPPVPSKLGWDEAMAIIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI 420
 Db 361 PMLTFGVQPNPPVPSKLGWDEAMAIIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI 420
 Qy 421 APWKGLLEETAVNDLISPPSGAPLISALGDLGGTHADVTAVPSTIIFTSVPVFTTGTSDVY 480
 Db 421 APWKGLLEETAVNDLISPPSGAPLISALGDLGGTHADVTAVPSTIIFTSVPVFTTGTSDVY 480
 Qy 481 AELNPSIIVRAGSPDPSSQPNDRHVAIFSTDDGKNWFGQSEPGVTTGGTVAASADGSRFV 540
 Db 481 AELNPSIIVRAGSPDPSSQPNDRHVAIFSTDDGKNWFGQSEPGVTTGGTVAASADGSRFV 540
 Qy 541 WAPGDPQPVVYAVFGNSWAASQGVPANAIQIRSDRVNPKTFYALSNGTFYRSTDDGVTF 600
 Db 541 WAPGDPQPVVYAVFGNSWAASQGVPANAIQIRSDRVNPKTFYALSNGTFYRSTDDGVTF 600
 Qy 601 QPVAAGLPSSGAVGMFHAVPGKEGDLWLAASSGLYHSTNGSSWSAITGVSSAVNVGFG 660
 Db 601 QPVAAGLPSSGAVGMFHAVPGKEGDLWLAASSGLYHSTNGSSWSAITGVSSAVNVGFG 660
 Qy 661 KSPAGSSYPVAVVGTGGTGYRSDDCGTTWVLIINDDOHQYGNWGOAITGDHANLRV 720
 Db 661 KSPAGSSYPVAVVGTGGTGYRSDDCGTTWVLIINDDOHQYGNWGOAITGDHANLRV 720
 Qy 721 YIGTNGRGIVYGDIGGAPSG 740
 Db 721 YIGTNGRGIVYGDIGGAPSG 740

RESULT 3

US-10-155-400-3
 ; Sequence 3, Application US/10155400
 ; Publication No. US20030108988A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DING, SHI-YOU
 ; APPLICANT: ADNEY, WILLIAM S.
 ; APPLICANT: VINZANT, TODD B.
 ; APPLICANT: HIMMEL, MICHAEL E.
 ; TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
 ; TITLE OF INVENTION: CELLULOLYTICUS
 ; FILE REFERENCE: NREL 01-36A
 ; CURRENT APPLICATION NUMBER: US/10/155,400
 ; CURRENT FILING DATE: 2002-10-22
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 740
 ; TYPE: PRT
 ; ORGANISM: Acidothermus cellulolyticus
 ; FEATURE: Catalytic domain GH74
 ; OTHER INFORMATION: Catalytic domain GH74
 US-10-155-400-3

Query Match 100.0%; Score 4036; DB 4; Length 740;
 Best Local Similarity 100.0%; Pred. No. 4.7e-284;
 Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ATTPYTWNSNVAIGGGFVDGIVFNEGAPGILYVRTDIGMYRWDAAANGRWIPLLDVWG 60
 Db 1 ATTPYTWNSNVAIGGGFVDGIVFNEGAPGILYVRTDIGMYRWDAAANGRWIPLLDVWG 60
 Qy 61 NNWYNGVWSIAADPINTNKVAAVGMVTSNDPNDGAILRSSDQATWQITPLPFLKGG 120
 Db 61 NNWYNGVWSIAADPINTNKVAAVGMVTSNDPNDGAILRSSDQATWQITPLPFLKGG 120

QY 121 NMPGRGMRERLAVDPNNNDILYFGAPSGKGLWRSTDGATWSQMTNFPDVGTYIANPTDT 180
DB 121 NMPGRGMRERLAVDPNNNDILYFGAPSGKGLWRSTDGATWSQMTNFPDVGTYIANPTDT 180
QY 181 TGYQSDIQGVVWVAFDKSSSLGQASKTIIFGVADPNPNPFWMSRDGGATWQAVPGAPTGF 240
DB 181 TGYQSDIQGVVWVAFDKSSSLGQASKTIIFGVADPNPNPFWMSRDGGATWQAVPGAPTGF 240
QY 241 IPHKGVDPVNVHLYIATNTGGPYDSSGVDWKFVSVTSGTWTRISPVPSSTDANDYFGY 300
DB 241 IPHKGVDPVNVHLYIATNTGGPYDSSGVDWKFVSVTSGTWTRISPVPSSTDANDYFGY 300
QY 301 SGLTIDROHPNTIMVATQISWMPDTIIFRSTDDGATWTRIDWTSYPNRSLRYVLDISAE 360
DB 301 SGLTIDROHPNTIMVATQISWMPDTIIFRSTDDGATWTRIDWTSYPNRSLRYVLDISAE 360
QY 361 PMLTFGVQPNPVPSPKLGWDEAMADPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI 420
DB 361 PMLTFGVQPNPVPSPKLGWDEAMADPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI 420
QY 421 APWKVGLSETAVNDLISPPSGAPLISALGDLGGFTHADVAVPSTIIFTSVFTTGTSDY 480
DB 421 APWKVGLSETAVNDLISPPSGAPLISALGDLGGFTHADVAVPSTIIFTSVFTTGTSDY 480
QY 481 AELNPSIIVRAGSDPSSQPNDRHVAFTDGGKWFQSGEPCGVTGTTVAASADGSRFV 540
DB 481 AELNPSIIVRAGSDPSSQPNDRHVAFTDGGKWFQSGEPCGVTGTTVAASADGSRFV 540
QY 541 WAPGDPGPVVYAVGFGNSWAASQGVANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF 600
DB 541 WAPGDPGPVVYAVGFGNSWAASQGVANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF 600
QY 601 QPVAAGLPSSGAVGVMFHAPCKEGDLWLAASSGLYHSTNGSSWSAITGYSSAVNVGFG 660
DB 601 QPVAAGLPSSGAVGVMFHAPCKEGDLWLAASSGLYHSTNGSSWSAITGYSSAVNVGFG 660
QY 661 KSAFGSSYPVAVVGTIGGVTGAYRSDCGTTWVLINDDQHOYGNWQAIGDHANLRV 720
DB 661 KSAFGSSYPVAVVGTIGGVTGAYRSDCGTTWVLINDDQHOYGNWQAIGDHANLRV 720
QY 721 YIGTNGRGIVYDGGAPSG 740
DB 721 YIGTNGRGIVYDGGAPSG 740

RESULT 4

US-10-155-400-6

; Sequence 6, Application US/10155400

; Publication No. US20030108988A1

; GENERAL INFORMATION:

; APPLICANT: DING, SHI-YOU

; APPLICANT: ADNEY, WILLIAM S.

; APPLICANT: VINZANT, TODD B.

; APPLICANT: HIMMEL, MICHAEL E.

; TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS

; TITLE OF INVENTION: CELLULOYTICUS

; FILE REFERENCE: NREL 01-36A

; CURRENT APPLICATION NUMBER: US/10/155,400

; CURRENT FILING DATE: 2002-10-22

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 740

; TYPE: PRT

; ORGANISM: Acidothermus cellulolyticus

US-10-155-400-6

Query Match 100.0%; Score 4036; DB 4; Length 740;

Best Local Similarity 100.0%; Pred. No. 4,7e-284;

Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATTPQYTWNSVAIGGGFVDGIVFNEGAPGILYVRTDIGMYRWDAAANGRWIPLLDWVGW 60

|||||

DB 1 ATTPQYTWNSVAIGGGFVDGIVFNEGAPGILYVRTDIGMYRWDAAANGRWIPLLDWVGW 60
QY 61 NNWNGVGVVSIADPINTNKVAAVCMYTNWSPNDGAILRSSDQAGATWQITPLPFLKGG 120
DB 61 NNWNGVGVVSIADPINTNKVAAVCMYTNWSPNDGAILRSSDQAGATWQITPLPFLKGG 120
QY 121 NMPGRGMRERLAVDPNNNDILYFGAPSGKGLWRSTDGATWSQMTNFPDVGTYIANPTDT 180
DB 121 NMPGRGMRERLAVDPNNNDILYFGAPSGKGLWRSTDGATWSQMTNFPDVGTYIANPTDT 180
QY 181 TGYQSDIQGVVWVAFDKSSSLGQASKTIIFGVADPNPNPFWMSRDGGATWQAVPGAPTGF 240
DB 181 TGYQSDIQGVVWVAFDKSSSLGQASKTIIFGVADPNPNPFWMSRDGGATWQAVPGAPTGF 240
QY 241 IPHKGVDPVNVHLYIATNTGGPYDSSGVDWKFVSVTSGTWTRISPVPSSTDANDYFGY 300
DB 241 IPHKGVDPVNVHLYIATNTGGPYDSSGVDWKFVSVTSGTWTRISPVPSSTDANDYFGY 300
QY 301 SGLTIDROHPNTIMVATQISWMPDTIIFRSTDDGATWTRIDWTSYPNRSLRYVLDISAE 360
DB 301 SGLTIDROHPNTIMVATQISWMPDTIIFRSTDDGATWTRIDWTSYPNRSLRYVLDISAE 360
QY 361 PMLTFGVQPNPVPSPKLGWDEAMADPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI 420
DB 361 PMLTFGVQPNPVPSPKLGWDEAMADPFNSDRMLYGTGATLYATNDLTKWDSGGQIHI 420
QY 421 APWKVGLSETAVNDLISPPSGAPLISALGDLGGFTHADVAVPSTIIFTSVFTTGTSDY 480
DB 421 APWKVGLSETAVNDLISPPSGAPLISALGDLGGFTHADVAVPSTIIFTSVFTTGTSDY 480
QY 481 AELNPSIIVRAGSDPSSQPNDRHVAFTDGGKWFQSGEPCGVTGTTVAASADGSRFV 540
DB 481 AELNPSIIVRAGSDPSSQPNDRHVAFTDGGKWFQSGEPCGVTGTTVAASADGSRFV 540
QY 541 WAPGDPGPVVYAVGFGNSWAASQGVANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF 600
DB 541 WAPGDPGPVVYAVGFGNSWAASQGVANAQIRSDRVNPKTFYALSNGTFYRSTDGGVTF 600
QY 601 QPVAAGLPSSGAVGVMFHAPCKEGDLWLAASSGLYHSTNGSSWSAITGYSSAVNVGFG 660
DB 601 QPVAAGLPSSGAVGVMFHAPCKEGDLWLAASSGLYHSTNGSSWSAITGYSSAVNVGFG 660
QY 661 KSAFGSSYPVAVVGTIGGVTGAYRSDCGTTWVLINDDQHOYGNWQAIGDHANLRV 720
DB 661 KSAFGSSYPVAVVGTIGGVTGAYRSDCGTTWVLINDDQHOYGNWQAIGDHANLRV 720
QY 721 YIGTNGRGIVYDGGAPSG 740
DB 721 YIGTNGRGIVYDGGAPSG 740

RESULT 5

US-09-917-376-1

; Sequence 1, Application US/09917376

; Publication No. US20040038334A1

; GENERAL INFORMATION:

; APPLICANT: DING, SHI-YOU

; APPLICANT: ADNEY, WILLIAM S.

; APPLICANT: VINZANT, TODD B.

; APPLICANT: HIMMEL, MICHAEL E.

; TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS

; TITLE OF INVENTION: CELLULOYTICUS

; FILE REFERENCE: 40197.4US01

; CURRENT APPLICATION NUMBER: US/09/917,376

; CURRENT FILING DATE: 2001-07-28

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 957

; TYPE: PRT

; ORGANISM: Acidothermus cellulolyticus

; FEATURE:

; NAME/KEY: MOD_RES

; LOCATION: (957)
; OTHER INFORMATION: Any amino acid
US-09-917-376-1

```

Query Match      100.0%; Score 4036; DB 3; Length 957;
Best Local Similarity 100.0%; Pred. No. 6.7e-284;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATTQPTWSNVAIGGGFVDGIVFNEGAPGILYVRTDIGMYRWDAAANGRWIPLLDWVGW 60
Db 47 ATTQPTWSNVAIGGGFVDGIVFNEGAPGILYVRTDIGMYRWDAAANGRWIPLLDWVGW 106
Qy 61 NNWGYNGVVSIAADPINTNKVAAVGMVYNSWDNDGAILRSSDQGATWQITPLPFKLG 120
Db 107 NNWGYNGVVSIAADPINTNKVAAVGMVYNSWDNDGAILRSSDQGATWQITPLPFKLG 166
Qy 121 NMPGRGGERLAVDPNNDNIIYFGAPSGKGLWRSTDGATWSQMTNFPDVGTYIANPTDT 180
Db 167 NMPGRGGERLAVDPNNDNIIYFGAPSGKGLWRSTDGATWSQMTNFPDVGTYIANPTDT 226
Qy 181 TGYQSDIQGVVWVAFDKSSSLGQASKTIIFGVADPNPNPVFWSRDGGATWQAVPGAPTGF 240
Db 227 TGYQSDIQGVVWVAFDKSSSLGQASKTIIFGVADPNPNPVFWSRDGGATWQAVPGAPTGF 286
Qy 241 IPHKGVPDPVNHVLIYATSNITGPGYDSSGVDWKFVSFTSGTWTTRISPVSTDTANDYFGY 300
Db 287 IPHKGVPDPVNHVLIYATSNITGPGYDSSGVDWKFVSFTSGTWTTRISPVSTDTANDYFGY 346
Qy 301 SGLTIDRQHPTIMVATOISWMPDTIIFRSTDGATWTRIWDWTSYPNRSRYVLDISAE 360
Db 347 SGLTIDRQHPTIMVATOISWMPDTIIFRSTDGATWTRIWDWTSYPNRSRYVLDISAE 406
Qy 361 PWLTFGVQPNPVPSPKLGWDEAMADPPNSDRMLYGTGATLYATNDLTKWDSGGQIHI 420
Db 407 PWLTFGVQPNPVPSPKLGWDEAMADPPNSDRMLYGTGATLYATNDLTKWDSGGQIHI 466
Qy 421 APVKGLEETAANDLISPPSGAPLISALGDLGGFTHADVTAVPSTIIFTSVFTTGTSDY 480
Db 467 APVKGLEETAANDLISPPSGAPLISALGDLGGFTHADVTAVPSTIIFTSVFTTGTSDY 526
Qy 481 AELNPSIIVRAGSPDPSSQPNDRHVAFTDGGKWFQSGEPGGVTTGGTVAASADGSRFV 540
Db 527 AELNPSIIVRAGSPDPSSQPNDRHVAFTDGGKWFQSGEPGGVTTGGTVAASADGSRFV 586
Qy 541 WAPGDPQPVVYAVFGNSWAASQGVPAQAIRSDRVNPKTFYALSNGTFFRSTDDGGVTF 600
Db 587 WAPGDPQPVVYAVFGNSWAASQGVPAQAIRSDRVNPKTFYALSNGTFFRSTDDGGVTF 646
Qy 601 QPVAAGLPSSGAVGMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG 660
Db 647 QPVAAGLPSSGAVGMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG 706
Qy 661 KSAPGSSYPAVFVGTGGTGVAYRSDCGTTWVLINDDQHOYGNWGOAITGDHANLRV 720
Db 707 KSAPGSSYPAVFVGTGGTGVAYRSDCGTTWVLINDDQHOYGNWGOAITGDHANLRV 766
Qy 721 YIGTNGRGIVYGDIGGAPSG 740
Db 767 YIGTNGRGIVYGDIGGAPSG 786

```

RESULT 6

US-10-155-400-1
; Sequence 1, Application US/10155400
; Publication No. US20030108988A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
; FILE REFERENCE: NREL 01-36A

; CURRENT APPLICATION NUMBER: US/10/155,400
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 957
; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (957)
; OTHER INFORMATION: Any amino acid
US-10-155-400-1

```

Query Match      100.0%; Score 4036; DB 4; Length 957;
Best Local Similarity 100.0%; Pred. No. 6.7e-284;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATTQPTWSNVAIGGGFVDGIVFNEGAPGILYVRTDIGMYRWDAAANGRWIPLLDWVGW 60
Db 47 ATTQPTWSNVAIGGGFVDGIVFNEGAPGILYVRTDIGMYRWDAAANGRWIPLLDWVGW 106
Qy 61 NNWGYNGVVSIAADPINTNKVAAVGMVYNSWDNDGAILRSSDQGATWQITPLPFKLG 120
Db 107 NNWGYNGVVSIAADPINTNKVAAVGMVYNSWDNDGAILRSSDQGATWQITPLPFKLG 166
Qy 121 NMPGRGGERLAVDPNNDNIIYFGAPSGKGLWRSTDGATWSQMTNFPDVGTYIANPTDT 180
Db 167 NMPGRGGERLAVDPNNDNIIYFGAPSGKGLWRSTDGATWSQMTNFPDVGTYIANPTDT 226
Qy 181 TGYQSDIQGVVWVAFDKSSSLGQASKTIIFGVADPNPNPVFWSRDGGATWQAVPGAPTGF 240
Db 227 TGYQSDIQGVVWVAFDKSSSLGQASKTIIFGVADPNPNPVFWSRDGGATWQAVPGAPTGF 286
Qy 241 IPHKGVPDPVNHVLIYATSNITGPGYDSSGVDWKFVSFTSGTWTTRISPVSTDTANDYFGY 300
Db 287 IPHKGVPDPVNHVLIYATSNITGPGYDSSGVDWKFVSFTSGTWTTRISPVSTDTANDYFGY 346
Qy 301 SGLTIDRQHPTIMVATOISWMPDTIIFRSTDGATWTRIWDWTSYPNRSRYVLDISAE 360
Db 347 SGLTIDRQHPTIMVATOISWMPDTIIFRSTDGATWTRIWDWTSYPNRSRYVLDISAE 406
Qy 361 PWLTFGVQPNPVPSPKLGWDEAMADPPNSDRMLYGTGATLYATNDLTKWDSGGQIHI 420
Db 407 PWLTFGVQPNPVPSPKLGWDEAMADPPNSDRMLYGTGATLYATNDLTKWDSGGQIHI 466
Qy 421 APVKGLEETAANDLISPPSGAPLISALGDLGGFTHADVTAVPSTIIFTSVFTTGTSDY 480
Db 467 APVKGLEETAANDLISPPSGAPLISALGDLGGFTHADVTAVPSTIIFTSVFTTGTSDY 526
Qy 481 AELNPSIIVRAGSPDPSSQPNDRHVAFTDGGKWFQSGEPGGVTTGGTVAASADGSRFV 540
Db 527 AELNPSIIVRAGSPDPSSQPNDRHVAFTDGGKWFQSGEPGGVTTGGTVAASADGSRFV 586
Qy 541 WAPGDPQPVVYAVFGNSWAASQGVPAQAIRSDRVNPKTFYALSNGTFFRSTDDGGVTF 600
Db 587 WAPGDPQPVVYAVFGNSWAASQGVPAQAIRSDRVNPKTFYALSNGTFFRSTDDGGVTF 646
Qy 601 QPVAAGLPSSGAVGMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG 660
Db 647 QPVAAGLPSSGAVGMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAITGVSSAVNVGFG 706
Qy 661 KSAPGSSYPAVFVGTGGTGVAYRSDCGTTWVLINDDQHOYGNWGOAITGDHANLRV 720
Db 707 KSAPGSSYPAVFVGTGGTGVAYRSDCGTTWVLINDDQHOYGNWGOAITGDHANLRV 766
Qy 721 YIGTNGRGIVYGDIGGAPSG 740
Db 767 YIGTNGRGIVYGDIGGAPSG 786

```

RESULT 7

US-10-156-761-9395

Db 524 DGDVLLMSSTSGALVSKSQG--TLTAVSSLPSCGAVIASDKSDNTVYFGSAGAIYVSK 580
Qy 595 DGGVTFQFVAAAGLPSSGAVGMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAI-TGVSS 653
Db 581 NTATSFKTIVS-LGSSITVNAI-RAHPSIAGDVWASTDKGLWHSTDYGSTFTQIGSGVTA 638
Qy 654 AVNVGFGKAPGSSYPVAVVGTIGGTGAYRSDDCGTTWVLINDDQHOYGNWQOA--IT 711
Db 639 GWSFGFGKASSTGSYVVIYGFITDGAAGLFKSEDACTNMQVSDASHGFGS-GSANVYN 697
Qy 712 GDHANLRVYIGTNGRGIVYGDIGGAPSG 740
Db 698 GDLQTYGRVFRGHERPGHLLRQSQREPAG 726

RESULT 9
US-10-155-400-7
; Sequence 7, Application US/10155400
; Publication No. US20030108988A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
; FILE OF INVENTION: CELLULOXYTICUS
; CURRENT APPLICATION NUMBER: US/01-36A
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 7
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Aspergillus aculeatus
US-10-155-400-7

Query Match 41.6%; Score 1680; DB 4; Length 726;
Best Local Similarity 46.2%; Pred. No. 4.7e-113;
Matches 346; Conservative 113; Mismatches 258; Indels 32; Gaps 17;
Qy 1 ATTQPYTWSNAI-GGGGFDGIVFNGAGPILYVRTDIGMYRWDANGRWIPLLDVWG 59
Db 1 AASQAYTWKVVVGGGGGFTPGIVFNPISAKVAVARTDIGAYRLN-SDDTWTFLMDVWG 59
Qy 60 ---WNNWGYNGVSIADPINTKNKVAAGMYTNSWDPNDGAILRSSDQATWQITPLPP 116
Db 60 NDTWHDW---GIDALATDPVTDTRVYVAVGMYTNEWDPNYSILRSTDDQGTWTKLPP 116
Qy 117 KLGNNPGRGWERLAVDPNNNDNLYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIAN 176
Db 117 KVGNNPGRGWERLAVDPNNKSNLYFGARSGHGLWKSTDYATGATWSNTFTWTGTYFQD 176
Qy 177 PTDITGQSDIQGVVVAFAFKSSSLQASAKTIFVGADPNPNPVFWSRDGATWQAQVGA 236
Db 177 SSST--YTSDPVGLAWVTFDSTSGSSGATPRIFGVADAGKSVFKSEDAATWAWVSGE 234
Qy 237 PT-GFPHKGVDPVNVHLVIAATNTGPGVGGSDGVWKFVSVTGWTTRISVPVSTDTAN 295
Db 235 PQYGFPLPHKGLVSPPEKTLVSIYANGAGPYDGTNGTVHKYNIITSGVWTDISP---TSLAS 291
Qy 296 DYFGYGLTITDRQHPNTIMVATQISWPDITIIIRSTDDGATWTRIMDWTSPNRSRYVL 355
Db 292 TYTGYGGLSVLDLQVPGTLVMAALNCWPPDELIIRSTDSGATWSPIWENWYPSINYYS 351
Qy 356 DISAEPWLTFGVQPNP-PVPSPLKGMDEAMADPPNSDRMLYGTGATLYATNDLTWQDS 414
Db 352 DISNAPWIDTSTDDQFPV---RVGWWVAALADPPDSNHLWLYGTGLTVYGGHDLTNWDS 408
Qy 415 GGOIHIAPMVKLEETAANDLISPPSGAPLISALGDLGGFTHADVAVPSTIFTSPVFTT 474
Db 409 KHNVTVKSLAVTEEMAVLGLITPPGFPALLSAVDDGGFVHSDLDAAPNQAYHTPYGT 468

Qy 475 GTSVDYAEINPSIIVRAGSFPDSSQPNDRHVAFTDGGKKNWFGSGEPGGVTTGCTVAASA 534
Db 469 TNGDITAGNKPNSIVRSASD--DYPT---LALSSNFGSTWYADYAASTSTGTGAVALSA 523
Qy 535 DGSFVWAPGDPGPVVVAVGVGNSWAAASQGPANAQIBSDRVNPKTFYALNNGTFVRS 594
Db 524 DGDVLLMSSTSGALVSKSQG---TLTAVSSLPSCGAVIASDKSDNTVYFGSAGAIYVSK 580
Qy 595 DGGVTFQFVAAAGLPSSGAVGMFHAVPGKEGDLWLAASSGLYHSTNGGSSWSAI-TGVSS 653
Db 581 NTATSFKTIVS-LGSSITVNAI-RAHPSIAGDVWASTDKGLWHSTDYGSTFTQIGSGVTA 638
Qy 654 AVNVGFGKAPGSSYPVAVVGTIGGTGAYRSDDCGTTWVLINDDQHOYGNWQOA--IT 711
Db 639 GWSFGFGKASSTGSYVVIYGFITDGAAGLFKSEDACTNMQVSDASHGFGS-GSANVYN 697
Qy 712 GDHANLRVYIGTNGRGIVYGDIGGAPSG 740
Db 698 GDLQTYGRVFRGHERPGHLLRQSQREPAG 726

RESULT 10
US-10-420-191-2
; Sequence 2, Application US/10420191
; Publication No. US20040067569A1
; GENERAL INFORMATION:
; APPLICANT: No. US20040067569A1ozymes Biotech, Inc.
; APPLICANT: Rey, Michael W.
; APPLICANT: Zaretsky, Elizabeth J.
; APPLICANT: Haas, Jeffrey A.
; TITLE OF INVENTION: Polypeptides Having Xyloglucanase Activity And Nucleic Acids
; FILE OF INVENTION: Encoding Same
; FILE REFERENCE: 10210-200-US
; CURRENT APPLICATION NUMBER: US/10/420,191
; CURRENT FILING DATE: 2003-04-18
; PRIOR APPLICATION NUMBER: US 60/373,987
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2
; LENGTH: 838
; TYPE: PRT
; ORGANISM: Trichoderma reesei
US-10-420-191-2

Query Match 39.1%; Score 1579; DB 4; Length 838;
Best Local Similarity 43.0%; Pred. No. 1.2e-105;
Matches 325; Conservative 123; Mismatches 264; Indels 44; Gaps 20;
Qy 6 YTWNSVAI-GGGGFDGIVFNGAGPILYVRTDIGMYRWDANGRWIPLLDVW---GW 60
Db 21 PSKNNKLVGGGGGVPGLIIFHFKTKGVAYARTDIGGLYRLN-ADDSWTAVTDGIADNAG 79

Qy 61 NNWGYNGVSIADPINTKNKVAAGMYTNSWDPNDGAILRSSDQATWQITPLPPKLG 120
Db 80 HNW---GIDAVALLPQDDQKVAAVGMVNSWDPNSGAILRSSDRGATWSFNLPKVVG 136
Qy 121 NMPGRGWERLAVDPNNNDNLYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTT 180
Db 137 NMPGRGAGERLAVDPANSNIIYFGARSGNGLWKSTDGGVTFPSKVSFTATGTYIPDPS 196
Qy 181 TGYQSDTQGVVWVAFDKSSSLQASAKTIFVGADPN-NNPVFWSRDGGATWQAVPCAPT 239
Db 197 NGYNSDKQGLMMVTFDSTSTTGGATSRIFVGTADNITASVYVSTNAGTSVAVPQPGK 256
Qy 240 FPHKGVDPVNVHLYIATSNITGPGYDGSVDVWKFVSVTGWTTRISVPVSTDTANDYFG 299
Db 257 YFPHKAKLPQAEKALYLTYSCTGYDGTGLSGVWRIDTAGGFWKDOIITVSGSDL---YFG 313
Qy 300 YSLGTLDRQHENTIMVATQISWPDITIIIRSTDDGATWTRIMDWTSPNRSRYVLDISA 359
Db 314 FGLGLDLQKQEGTLVAVSLNSWMPDAQLFRSTDSCTTWSPIWAWASYPTETYYISISTPK 373

QY 360 EPWLTFG-VQPNPPVPS-----PKLGMDEAMADPPNSDRMLYGTGATLYATNDLTKWDS 414
 DB 374 APWIKNNFIDVTSSEPSDGLIKRLGWMIESLEIDPTDSNHWLYGTGMTIFGGHDLTNWDT 433
 QY 415 GGOIHAPMWKLEETAVNDLISPPSGAPLISALGDLGGFTHA---DVTAVPSTIFTSPV 471
 DB 434 RHNVISQSLADGIEBFSVQDLASAPGSELLAAVGDNGFTFASRNDLGTSPQTVAWATPT 493
 QY 472 FTTGTSVDYAEINLPSIIIVRAGSFDPSSQPNDRHVAFPSTDGKKNWFGQSEPGGVTTGGTVA 531
 DB 494 WATSTSDYAGNSVKSVVRVGN---VAISSDGGATWSIDYAADTSMNGGTVA 548
 QY 532 ASADGSRFWAPCDPQGPVYVAVGFGNSWAASQGVPAQAIRSDRVNPKTFYALNSGTFFY 591
 DB 549 YSADGDTILWSTASSG---VQSRQFQGSFASVSSLPAGAVIASDKKTNVSFVYAGSGSTFY 605
 QY 592 RSTGCGVTFQPVAAAGLPSSGAVGVN--FHAVPGKEGDLWLAASSGLYHSTNGGSSWSAI- 648
 DB 606 VSKDTGSSF---TRG-PKLGSGAGTIRDIAAHPPTAGTLVSTDVGFIRSTDSTGTTFGQVS 661
 QY 649 TGVSANVNVGFGKSAPGSSYPVAVFVGTIGGVTGA--YRSDCCTGTTWVLINDDQHQYGNW 706
 DB 662 TALTNTYQIALGVGS-GSNW-NLYAFGT--GPSGARLYASGDSGASWTDIQSQGFGSID 717
 QY 707 GOAITGDHANLRVYIGTNGRIVY--GDIGGAPSG 740
 DB 718 STKVAGSGSTAGQVYVGTNGRGRVFAQGTGGGTGG 753

RESULT 11

US-10-026-994-2
 ; Sequence 2, Application US/10026994
 ; Publication No. US20030113732A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dunn-Coleman, Nigel
 ; APPLICANT: Goedegebuur, Frits
 ; APPLICANT: Ward, Michael
 ; APPLICANT: Yao, Jian
 ; TITLE OF INVENTION: EGVI Endoglucanase and Nucleic Acids
 ; TITLE OF INVENTION: Encoding the Same
 ; FILE REFERENCE: GC698
 ; CURRENT APPLICATION NUMBER: US/10/026,994
 ; PRIOR FILING DATE: 2002-04-30
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 818
 ; TYPE: PRT
 ; ORGANISM: Trichoderma reesei
 US-10-026-994-2

Query Match 39.0%; Score 1573.5; DB 4; Length 818;
 Best Local Similarity 42.9%; Pred. No. 3e+105;
 Matches 324; Conservative 122; Mismatches 265; Indels 45; Gaps 20;

QY 6 YTWNSVAI-GGGGFGVDGIVNEGAPGILYVTRDTCGMRYMDAANGRWIPLLDWV---CW 60
 DB 2 FSKNVKLVGGGGFVPGFIIFPKTKGVAYARTDTCGLYRLN-ADDSWTAVTDTGIADNAGW 60
 QY 61 NNWNGYVVVIAADPINTNKVAAVGMVYTNSDPNDGAILRSSDQATWQITPLPKLGG 120
 DB 61 HNW---GIDVALDPQDQKVAAVGMVYTNSDPNSGAILIRSSDRGATWSFNLNPKYVG 117
 QY 121 NMPGRGGERLAVDPNNNLIYFGAPSGKGLWRSTDSGATWSQMTNFFDVGYIANTPTD 180
 DB 118 NMPGRGGERLAVDPANSNLIYFGARSGNLWKSTDGCVTFKSVSFTATGYIIPDPSDS 177
 QY 181 TGYQSDIQGVVWAFDKSSSLGQASKTIFVGVADP-NNPVFWSRDGGATWAOVAPGTG 239
 DB 178 NGYNSDKGLMWVTFDSTSSSTGGATSRIFVGTADNITASVYVSTNAGSTWSAVFCQPGK 237
 QY 240 FIPHGKVPDPNVHLYIATSNITGGPYDGGSDGVKFSVTSGTWTIRISVPVSTDTANDYFG 299

DB 238 YPHKAKLOPAEKALYLTYSDCGTGPDYDGTILGSVVRVDIAGGTWKDITPVSGSDL---YFG 294
 QY 300 YEGLTIDRQHPNTIMVATOISWMPDTIIFRSTDGGATWTRIDWTSYPNRSRLRYVLDISA 359
 DB 295 FGLGLDLQKPGTLVVASLNSWMPDAQLFRSTDSGTGTSPIWAWASYPTETYYYSITPK 354
 QY 360 EPWLTFG-VQPNPPVPS-----PKLGMDEAMADPPNSDRMLYGTGATLYATNDLTKWDS 414
 DB 355 APWIKNNFIDVTSSEPSDGLIKRLGWMIESLEIDPTDSNHWLYGTGMTIFGGHDLTNWDT 414
 QY 415 GGOIHAPMWKLEETAVNDLISPPSGAPLISALGDLGGFTHA---DVTAVPSTIFTSPV 471
 DB 415 RHNVISQSLADGIEBFSVQDLASAPGSELLAAVGDNGFTFASRNDLGTSPQTVAWATPT 474
 QY 472 FTTGTSVDYAEINLPSIIIVRAGSFDPSSQPNDRHVAFPSTDGKKNWFGQSEPGGVTTGGTVA 531
 DB 475 WATSTSDYAGNSVKSVVRVGN---TAGTQVAISSDGGATWSIDYAADTSMNGGTVA 528
 QY 532 ASADGSRFWAPCDPQGPVYVAVGFGNSWAASQGVPAQAIRSDRVNPKTFYALNSGTFFY 591
 DB 529 YSADGDTILWSTASSG---VQSRQFQGSFASVSSLPAGAVIASDKKTNVSFVYAGSGSTFY 585
 QY 592 RSTGCGVTFQPVAAAGLPSSGAVGVN--FHAVPGKEGDLWLAASSGLYHSTNGGSSWSAI- 648
 DB 586 VSKDTGSSF---TRG-PKLGSGAGTIRDIAAHPPTAGTLVSTDVGFIRSTDSTGTTFGQVS 641
 QY 649 TGVSANVNVGFGKSAPGSSYPVAVFVGTIGGVTGA--YRSDCCTGTTWVLINDDQHQYGNW 706
 DB 642 TALTNTYQIALGVGS-GSNW-NLYAFGT--GPSGARLYASGDSGASWTDIQSQGFGSID 697
 QY 707 GOAITGDHANLRVYIGTNGRIVY--GDIGGAPSG 740
 DB 698 STKVAGSGSTAGQVYVGTNGRGRVFAQGTGGGTGG 733

RESULT 12

US-10-156-761-10111
 ; Sequence 10111, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; PRIOR FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 10111
 ; LENGTH: 739
 ; TYPE: PRT
 ; ORGANISM: Streptomyces avermitilis
 US-10-156-761-10111

Query Match 34.8%; Score 1405; DB 4; Length 739;
 Best Local Similarity 40.8%; Pred. No. 4.4e-93;
 Matches 308; Conservative 104; Mismatches 257; Indels 86; Gaps 21;

QY 6 YTWNSVAI-GGGGFGVDGIVNEGAPGILYVTRDTCGMRYMDAANGRWIPLLDWGNNWGY 65
 DB 42 YRWRNAVIGTGFVTVGLFHPHSVRGLAYARTDTCGAYRWDGDRGARWTPLDLHLGWDWNL 101
 QY 66 NGVWSTAADPINTNKVAAVGMVYTNSDPNDGAILRSSDQATWQITPLPKLGGNNMGR 125
 DB 102 LGVEAMAVDTPHDLRLYLAVGYAQSWAGN-GAVLRSDEGATWTRTDLTVKLGGNEDGR 160

QY 126 GMERLAVDPNNNIIYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQS 185
 Db 161 GAGERLLVDPDSDTLWLGT-RHDGLLKSTDRGATMAAATAP-----AKANS 207
 QY 186 DIOGVVVAADFSSSISGQASKTIFVGVADPNP-----VWSDRGATQAVGAPGTF 240
 Db 208 SGOGVVF-----LVAAGRTVYAGWGDGSGTANLYRTAD-GTTWGAVPGRPSGT 257
 QY 241 ---IPHKGVDPVNVHLYIATSTNGPYDGSDDVWKFSTVGTWTRISPV-----PSTDT 293
 Db 258 SAKVPLRAAYDTHRELYVYTGAPGCGSDSVHKLRTATGTWTEVTPVKCGTISDG 317
 QY 294 ANDYFGSLTIDRQHPNTIMVATQISWMPDTTIFRSTDGATWTRIDWTSYPNRSLRY 353
 Db 318 SADTFAYGGVAVDARPGTLVSTNNRWADGDTVFRSTDGRTWTSKLD-----AA 368
 QY 354 VLDISAPMLTFFGQVQNPVPSKLGWDEAMALDPNSDRMLYGTGATLYATNDLTWTD 413
 Db 369 VFDVSETPFLDWGDD-----KPKFGWMIQALAVDPYDSQHVYVGTGATLYGTRDLKRW- 421
 QY 414 SGQOIHIAPMVKGLERTAVNDLISPPSG-APLISALGDLGFTHADVTAVPST-IFTSPV 471
 Db 422 -----APRILEESAVRQLISPPVGEAHLISGLDIGNVHERLUTASPSGMAITNPV 474
 QY 472 FTTGTSVDVAELNPSIIVRAGSPDPSPQNDRHVAFSTDGKKNW--PQSEPGGVTTGGT 529
 Db 475 FGSATGLAQAAAPAYVVRTGWGDHNG-----AYSHDGRTWAPFEAPQADKAPG 528
 QY 530 VAASAGSRFWA-----PDGPGPVVA-----VGFNGSWAASQGVANAIQIRSDRVNPKT 581
 Db 529 IATSADGGTLLMSFVHWG-----TTVAHRSTDNGASWSEVSFPKGAFTPADPADPTR 583
 QY 582 FYA--LSNGTFFYSTDGGVTTFQVAAGLPSSGAVGVMFHAPVCKEGDLWLAAS-SGLYHS 638
 Db 584 FYAYDFNGLYASTDGSRTARAGLP-SGDSQFKLVAAPGRSGDLWLSAKWNGLYRS 642
 QY 639 TNGSSSAITGVSSAVNVGFGKSPAGSSYPVAVFVGTIGVGTGAYRSDDCGTTWVLND 698
 Db 643 TDGDTFARIDSCWASYTLGFGAAGADYPAIYQVGSSTETITAVYRSDDAARTWRLND 702
 QY 699 DHOYQNWGAITGDHANLRVYVGTNGRGIVYGD 733
 Db 703 DAHOWGICEAVVGDPRHGRVYLATNGRGIQYE 737
 RESULT 13
 US-10-395-241-14
 ; Sequence 14, Application US/10395241
 ; Publication No. US20040038367A1
 ; GENERAL INFORMATION:
 ; APPLICANT: YAOI, Katsuro
 ; APPLICANT: MITSUISHI, Yasushi
 ; TITLE OF INVENTION: NOVEL XYLOGLUCAN OLIGOSACCHARIDE-DEGRADING ENZYME, POLYNUCLEOTIDE
 ; FILE REFERENCE: 073756
 ; CURRENT APPLICATION NUMBER: US/10/395,241
 ; PRIOR FILING DATE: 2003-03-25
 ; PRIOR APPLICATION NUMBER: JP 2002-83433
 ; PRIOR FILING DATE: 2002-03-25
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 14
 ; LENGTH: 789
 ; TYPE: PRT
 ; ORGANISM: Geotrichum sp. M128
 US-10-395-241-14
 Query Match 27.7%; Score 1117; DB 4; Length 789;
 Best Local Similarity 35.0%; Pred. No. 3.9e-72;
 Matches 281; Conservative 124; Mismatches 279; Indels 120; Gaps 31;
 QY 6 YTWNSVAIGGGVFDGIVFNEGAPGILYVTDIGGMVWDAANGRWIPLLDWGMNNGY 65

Db 4 YEFKNVAIGGGYITGIVAHPKTKDLYARTDIGAYRWDACTSKWIPLNDFIEAQMNI 63
 QY 66 NGVYSIAADPINTKNWAAAGMYT-NSWDPNDDGALLRSSDQCATWQITPLPKLGNMGP 124
 Db 64 MGTESIALDPNPRLYLAQRYGVDEW-----AAFVSEDRQCSFTIYESPPMGANDMG 119
 QY 125 RGMERLAVDPNNNIIYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTGYQ 184
 Db 120 RNNGERLAVNFNSNEVMMGTRT-EGIKSSDRAKTWTNVTISPDAF-----TNGIGYT 172
 QY 185 SDIOGVVVAADFSSSISGQASKTIFVGVADPNPVPFWSRGGATQAVGAPGTFI--- 241
 Db 173 S-----VIEDP-----ERNGTIYASATAPQG-MYVTHDGVSWEPVAGQPSSSLNRT 218
 QY 242 -----PHKGVDPVNVHLYIATSTNGPYDGSDDVWKFSTVGTWTRISPV----- 285
 Db 219 TGAPDPPKASIAPOPMKVALTP--NFLVTVADYPPGPGWVTFGEVWRQNRSTGAWDDIT 276
 QY 286 -----SPVPTDANDYFGYSLTIDRQHPNTIMVATQISWMPDTI---IFRSTDGAT 336
 Db 277 PRVGNSSPAPYNNQTFPAGGFCGLSDATNPNLVIT-LDRDPGALDSIYLSTDAAT 335
 QY 337 WTRIWDWTS-----YPNRSRLRYLDISAEPMLTFFGVQV-----NPPVPSP---KLG 379
 Db 336 WKDVTQLSSPSNLEGNWGHPTAARY-KDGTVPVWLDNFNNGPQWGGYGAPHGTPGLTKFG 394
 QY 380 WMEAMALDPNSDRMLYGTGATLYATNDLTG-----WDSGGQIHIAP-----MVKGLEETA 431
 Db 395 WMSAVLIDPFPBHLMYGTGATITWATDLSRVEKW-----APSWYLOIDGIEENA 446
 QY 432 VNDLISPPSGAPLISALGDLGFTHADVTAVPSTIFTSPVFTTGTSDVVAELNPSIIVRA 491
 Db 447 ILSRSPKSGAALLSGIDISGMKEDDLTK-PQMFAGAPQFNSNLSDIDAAGNFPNVVRA 505
 QY 492 GSPDPSPQNDRHVAFSTDGKKNW--FQSEBPGGVTT---GGTVAASADGSRFWAP--G 544
 Db 506 GSSGHEYDSACARGAYATDGDWATIFPTCPGPNWASHVQGSTIAVDASGSIWSTKLD 565
 QY 545 DPGQVTVYAVFGNSWAAASQGVPA-----NAQIRSDRVNPKTFYALNSGTFTYRSTDGV 598
 Db 566 EQASGPWYSHDYGKTWS-----VPAGDLKAQATANVLSKVDQGTFTYATDGKGFVSTDGK 621
 QY 599 TFQVAAAGLPSSGAVGVMFHAPVCKEGDLWL-AAASSGLYHSTNGGSSKSAI-TGVSSAVN 656
 Db 622 SYAAGKAGLV--GTSLMPAVNVPVAGVWVVPVEGGLFHSDFGASFTRVGTANATLVS 679
 QY 657 VGFGKS-----APGSSYPVAVFVVGTT--IGGVTCAYRSDDCGTTWVLINDDHOYQNWGA 710
 Db 680 VGAPKSKDGKASAPSAVFWGTDKPGSDIGLYRSDDNGSTWTRVNDQEHYSG-PTWI 738
 QY 711 TGDHANLRVYVGTNGRGIVYGD 734
 Db 739 EADPKVYGRVYLTNGRGIVYADL 762
 RESULT 14
 US-10-395-241-12
 ; Sequence 12, Application US/10395241
 ; Publication No. US20040038367A1
 ; GENERAL INFORMATION:
 ; APPLICANT: YAOI, Katsuro
 ; APPLICANT: MITSUISHI, Yasushi
 ; TITLE OF INVENTION: NOVEL XYLOGLUCAN OLIGOSACCHARIDE-DEGRADING ENZYME, POLYNUCLEOTIDE
 ; FILE REFERENCE: 073756
 ; CURRENT APPLICATION NUMBER: US/10/395,241
 ; CURRENT FILING DATE: 2003-03-25
 ; PRIOR APPLICATION NUMBER: JP 2002-83433
 ; PRIOR FILING DATE: 2002-03-25
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 12

; LENGTH: 812
; TYPE: PRT
; ORGANISM: Geotrichum sp. M128
US-10-395-241-12

Query Match 27.7%; Score 1117; DB 4; Length 812;
Best Local Similarity 35.0%; Pred. No. 4e-72;
Matches 281; Conservative 124; Mismatches 279; Indels 120; Gaps 31;

```
QY 6 YTNVAIGGGFVDGIVNEGAPGILYVRTDIGMYRMDAANGRWIPLLDWVGNWNGY 65
Db 27 YEFKNVAIGGGYITGIVAHPKTDLLYARTDIGAYRWDAGTSKWIPLNDPIEAQDMNI 86
QY 66 NGVSIADDPINTNKVMAAVGMYT-NSWDPNDGAILRSSDQATWQITPLPKLGNMGP 124
Db 87 MGTESIALDPNPDRLYLAAQRYVGDEW----AAFVSEDRGQSFTIYESPPMGANDMG 142
QY 125 RGMGERLAVDPNNDNLLYFGAPSGKGLWRSTDGATWSOMTNFPDVGTYIANPTDTTGYQ 184
Db 143 RNNGERLAVNPNFNSNEVWMTGRT-EGIKSSDRAKTWTNVTSPIDAF-----TNGIGYT 195
QY 185 SDIQGVVWVAFDKSSSLGQASKTIFVGVADPNPNPVFWSRDGATWQAVPGAPTGF1--- 241
Db 196 S-----VIFDP-----ERNGTIYASATAPOG-MYVTHDGGVSWEPVAGOPSSWLNRT 241
QY 242 -----PHKGVDPVNVHLYIATSNCTGPDYDSSGDVWKFVSVTSGTWTRI- 285
Db 242 TGAFPDKKPASTAPQPMKVALTP--NFLYVYADYFGPWGVTFGKWRQNRNRTSGAWDDIT 299
QY 286 -----SPVSTDTANDYFGYGLTIDROHPNTIMVATQISWMPDTI---IFRSTDGAT 336
Db 300 PRVGNSSPAPYNNOTFPAGGFCGLSVDATNPNRLVIT-LDRDPGPALEDSIYLSIDAGAT 358
QY 337 WTRIDWMTS-----YPNRSLRYVLDISAEPMLTFGVOP-----NPPVPSP---KLG 379
Db 359 WKDVTQLSSPSNLEGNWGHPTNAARY-KDGTVPVPLDFNNGPQWGGYGAPHGTPGLTKFG 417
QY 380 WMDEAMADPFNSDRMLYGTGATLYATNDLTK---WDSGGQIHIAP-----MVKGLEETA 431
Db 418 WMSAVLIDPFNPEHLMYGTGATITWATDLSRVEKDW-----APSWYLQIDGIEENA 469
QY 432 VNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSDVYAEINPSIIVRA 491
Db 470 ILSLRSPKSGAALLSGIGDISGMKHDDLTK-PQKMFAGAPQFSLNLSIDAAGNPNVAVRA 528
QY 492 GSFDPSSQPNDRHVAFAFSTDGKNW--FQSGEPGGVTT--GGTVAASADGSRFWAP--G 544
Db 529 GSSGHEYDSACARGAYATDGDGAWTTFPTCPPGMNASHYQGSTIAVDASSQIWMSTKLD 588
QY 545 DPGQPVVYAVFGNSWAASQGVPA-----NAQIRSDRVNPKTFYALNNGTFYRSTDGGV 598
Db 589 EQASGPWYSHDYGKTWS-----VPAGDLKAQTANVLSDKVQGTGYATDGGKFFVSTDGK 644
QY 599 TFQPVAAGLPSSGAVGVMFHAPVPGKEGLWL-AAASSGLYHSTNGGSSWSAI-TGVSSAVN 656
Db 645 SYAAKGAGLVT--GTSMPAVNPVWAGDVVVPVEGGLFHSITDFGASFTRVGTANATLVS 702
QY 657 VGFGKS-----APGSSYPAVFVVGTT--IGGVGTAYRSDDCGTTWVLINDDQHOYGNWQAI 710
Db 703 VGAPKSKSDGKKASAPSAVFITWGTDKPGSDIGLYRSDDNGSTWTRVNDQEHNSYSG-PTMI 761
QY 711 TGDHANLRVYVIGTNGRIGIVYDI 734
Db 762 EADPKVGRVVLGTNGRIGIVYADL 785
```

RESULT 15

US-10-395-241-18
; Sequence 18, Application US/10395241
; Publication No. US20040038367A1
; GENERAL INFORMATION:
; APPLICANT: YAOI, Katsuro
; APPLICANT: MITSUISHI, Yasushi

; TITLE OF INVENTION: NOVEL XYLOGLUCAN OLIGOSACCHARIDE-DEGRADING ENZYME, POLYNUCLEOTIDE
; FILE REFERENCE: Q73756
; CURRENT APPLICATION NUMBER: US/10/395,241
; PRIOR FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: JP 2002-83433
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 18
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Xyloglucan Oligosaccharide-Degrading Enzyme with Histidine Tag
US-10-395-241-18

Query Match 27.6%; Score 1114; DB 4; Length 826;

Best Local Similarity 35.0%; Pred. No. 6.8e-72;
Matches 281; Conservative 123; Mismatches 280; Indels 120; Gaps 31;

```
QY 6 YTNVAIGGGFVDGIVNEGAPGILYVRTDIGMYRMDAANGRWIPLLDWVGNWNGY 65
Db 5 YEFKNVAIGGGYITGIVAHPKTDLLYARTDIGAYRWDAGTSKWIPLNDPIEAQDMNI 64
QY 66 NGVSIADDPINTNKVMAAVGMYT-NSWDPNDGAILRSSDQATWQITPLPKLGNMGP 124
Db 65 MGTESIALDPNPDRLYLAAQRYVGDEW----AAFVSEDRGQSFTIYESPPMGANDMG 120
QY 125 RGMGERLAVDPNNDNLLYFGAPSGKGLWRSTDGATWSOMTNFPDVGTYIANPTDTTGYQ 184
Db 121 RNNGERLAVNPNFNSNEVWMTGRT-EGIKSSDRAKTWTNVTSPIDAF-----TNGIGYT 173
QY 185 SDIQGVVWVAFDKSSSLGQASKTIFVGVADPNPNPVFWSRDGATWQAVPGAPTGF1--- 241
Db 174 S-----VIFDP-----ERNGTIYASATAPOG-MYVTHDGGVSWEPVAGOPSSWLNRT 219
QY 242 -----PHKGVDPVNVHLYIATSNCTGPDYDSSGDVWKFVSVTSGTWTRI- 285
Db 220 TGAFPDKKPASTAPQPMKVALTP--NFLYVYADYFGPWGVTFGKWRQNRNRTSGAWDDIT 277
QY 286 -----SPVSTDTANDYFGYGLTIDROHPNTIMVATQISWMPDTI---IFRSTDGAT 336
Db 278 PRVGNSSPAPYNNOTFPAGGFCGLSVDATNPNRLVIT-LDRDPGPALEDSIYLSIDAGAT 336
QY 337 WTRIDWMTS-----YPNRSLRYVLDISAEPMLTFGVOP-----NPPVPSP---KLG 379
Db 337 WKDVTQLSSPSNLEGNWGHPTNAARY-KDGTVPVPLDFNNGPQWGGYGAPHGTPGLTKFG 395
QY 380 WMDEAMADPFNSDRMLYGTGATLYATNDLTK---WDSGGQIHIAP-----MVKGLEETA 431
Db 396 WMSAVLIDPFNPEHLMYGTGATITWATDLSRVEKDW-----APSWYLQIDGIEENA 447
QY 432 VNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSDVYAEINPSIIVRA 491
Db 448 ILSLRSPKSGAALLSGIGDISGMKHDDLTK-PQKMFAGAPQFSLNLSIDAAGNPNVAVRA 506
QY 492 GSFDPSSQPNDRHVAFAFSTDGKNW--FQSGEPGGVTT--GGTVAASADGSRFWAP--G 544
Db 507 GSSGHEYDSACARGAYATDGDGAWTTFPTCPPGMNASHYQGSTIAVDASSQIWMSTKLD 566
QY 545 DPGQPVVYAVFGNSWAASQGVPA-----NAQIRSDRVNPKTFYALNNGTFYRSTDGGV 598
Db 567 EQASGPWYSHDYGKTWS-----VPAGDLKAQTANVLSDKVQGTGYATDGGKFFVSTDGK 622
QY 599 TFQPVAAGLPSSGAVGVMFHAPVPGKEGLWL-AAASSGLYHSTNGGSSWSAI-TGVSSAVN 656
Db 623 SYAAKGAGLVT--GTSMPAVNPVWAGDVVVPVEGGLFHSITDFGASFTRVGTANATLVS 680
QY 657 VGFGKS-----APGSSYPAVFVVGTT--IGGVGTAYRSDDCGTTWVLINDDQHOYGNWQAI 710
Db 681 VGAPKSKSDGKKASAPSAVFITWGTDKPGSDIGLYRSDDNGSTWTRVNDQEHNSYSG-PTMI 739
```

QY 711 TGDHANLRRVYIGTNGRGIVYVDI 734
Db 740 EADPKVYGRVYLGTVNGRGIVYADL 763

Search completed: March 2, 2006, 14:26:54
Job time : 146.551 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 14:27:20 ; Search time 15.0053 Seconds
(without alignments)
986.322 Million cell updates/sec

Title: US-09-917-376-3

Perfect score: 4036

Sequence: 1 ATQPTWNSNAIGGGGFVD.....YIGTNGRGIVYGDIGGAPSG 740

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 135339 seqs, 20000136 residues

Total number of hits satisfying chosen parameters: 135339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	175.5	4.3	2523	7	US-11-052-554A-143	Sequence 143, App
2	168.5	4.2	5291	7	US-11-052-554A-281	Sequence 281, App
3	166.5	4.1	980	7	US-11-052-554A-17	Sequence 17, App
4	160.5	4.0	3716	7	US-11-052-554A-141	Sequence 141, App
5	156.5	3.9	2204	7	US-11-052-554A-134	Sequence 134, App
6	156	3.9	3157	7	US-11-052-554A-142	Sequence 142, App
7	155	3.8	7465	7	US-11-087-099-7521	Sequence 7521, App
8	153.5	3.8	807	7	US-11-087-099-12161	Sequence 12161, A
9	148.5	3.7	894	7	US-11-087-099-11214	Sequence 11214, A
10	147.5	3.7	755	6	US-10-517-939-330	Sequence 330, App
11	146	3.6	949	7	US-11-052-554A-6	Sequence 6, Appli
12	145.5	3.6	3300	7	US-11-052-554A-133	Sequence 133, App
13	145	3.6	760	7	US-11-052-554A-68	Sequence 68, Appl
14	140.5	3.5	955	7	US-11-052-554A-179	Sequence 179, App
15	140.5	3.5	1461	7	US-11-052-554A-283	Sequence 283, App
16	139	3.4	3063	7	US-11-186-284-26	Sequence 26, Appl
17	137.5	3.4	1436	7	US-11-052-554A-140	Sequence 140, App
18	137	3.4	1424	7	US-11-102-476-4	Sequence 4, Appli
19	136.5	3.4	393	5	US-09-910-864-1	Sequence 1, Appli
20	134.5	3.3	1007	6	US-10-517-939-84	Sequence 84, Appl
21	133.5	3.3	837	7	US-11-052-554A-159	Sequence 159, App
22	133.5	3.3	1889	7	US-11-102-476-46	Sequence 46, Appl
23	133.5	3.3	3132	7	US-11-087-099-1245	Sequence 1245, App
24	130	3.2	1588	7	US-11-052-554A-280	Sequence 280, App
25	129.5	3.2	736	7	US-11-085-185-4	Sequence 4, Appli

26	129	3.2	1306	7	US-11-052-554A-139	Sequence 139, App
27	126.5	3.1	778	7	US-11-052-554A-144	Sequence 144, App
28	126	3.1	527	7	US-10-886-504-10	Sequence 10, Appl
29	126	3.1	527	6	US-10-886-505-10	Sequence 10, Appl
30	126	3.1	527	6	US-10-886-527-10	Sequence 10, Appl
31	126	3.1	2105	7	US-11-052-554A-173	Sequence 173, App
32	125.5	3.1	538	7	US-11-174-398-16	Sequence 16, Appl
33	125.5	3.1	538	7	US-11-132-947-6	Sequence 6, Appli
34	125.5	3.1	538	7	US-11-197-488-2	Sequence 2, Appli
35	125	3.1	527	6	US-10-886-504-8	Sequence 8, Appli
36	125	3.1	527	6	US-10-886-505-8	Sequence 8, Appli
37	125	3.1	527	6	US-10-886-527-8	Sequence 8, Appli
38	125	3.1	611	6	US-10-517-939-156	Sequence 156, App
39	125	3.1	1417	7	US-11-052-554A-8	Sequence 8, Appli
40	125	3.1	6893	7	US-11-205-109-14	Sequence 14, Appl
41	124.5	3.1	461	7	US-11-087-099-11993	Sequence 11993, A
42	124.5	3.1	1579	7	US-11-052-554A-9	Sequence 9, Appli
43	124	3.1	525	7	US-11-098-686-10993	Sequence 10993, A
44	124	3.1	957	7	US-11-108-172-1065	Sequence 1065, Ap
45	122.5	3.0	863	7	US-11-087-099-10485	Sequence 10485, A

ALIGNMENTS

RESULT 1

US-11-052-554A-143

; Sequence 143, Application US/11052554A

; Publication No. US20050288866A1

; GENERAL INFORMATION:

; APPLICANT: Sachdeva, et al.

; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL

; FILE REFERENCE: 30853/40359A

; CURRENT APPLICATION NUMBER: US/11/052,554A

; CURRENT FILING DATE: 2005-02-07

; PRIOR APPLICATION NUMBER: US 60/589,227

; PRIOR FILING DATE: 2004-07-20

; PRIOR APPLICATION NUMBER: IN 173/DEL/2004

; PRIOR FILING DATE: 2004-02-06

; NUMBER OF SEQ ID NOS: 763

; SOFTWARE: Patent in version 3.3

; SEQ ID NO 143

; LENGTH: 2523

; TYPE: PRT

; ORGANISM: Mycobacterium tuberculosis H37Rv

US-11-052-554A-143

Query Match	4.3%;	Score 175.5;	DB 7;	Length 2523;
Best Local Similarity	21.8%;	Pred. No. 0.00041;		
Matches 175;	Conservative 92;	Mismatches 327;	Indels 209;	Gaps 38;
QY	10	NVAICGGGFGDGVFNEGA--PGILYVRDTCGMVYRWDAAN-GRWIPLLDVGWGNWYN	66	
Db	1522	NTGICNSGIASGTGLFNAGGFGTGVVAGSYNTGSPNAGQANTGGFNGSVNTGMLNTG--	1579	
QY	67	GVVSTAADPINTKVAAGVMTNSWDNDGAILRSS--DQATWQITPLPKLGGNMPG-	124	
Db	1580	-----DINTG-----VANSQDVNTGAFISGNYNGAFWR-----GDYQGL	1614	
QY	125	RMGERLAVDPNNDILYFGAPSGKGLWRSTDSGATWSQMTNPPDVGTYIANPTDITGYQ	184	
Db	1615	LGFSYRPAVLQ-----TPFLDLTLTGGJGSGVVIPAIDPAIR	1652	
QY	185	SDIQGVWVW-AFDKSSSSISLQ---ASKTIFVGVAD-----PNNPVFWSRGGATWQA	232	
Db	1653	PEFSANVAIDSTVSPIPQIDLAATVSVGLGPITVPHLDIPRPVTLNLYFGSQ---	1709	
QY	233	VGAFTGPIPHKGFPDPVNNHLYIATSNVTGGPYDSSGSDGVKPSVTSQWTRISPVSTD	292	
Db	1710	-PGGPKIGPTGLFNTPIGLTPALSQIVIGASSSQGITAFANLPLPFTPTVVTIDEIP	1768	
QY	293	TANDYFGYS-----GLTIDRQHPNTI-----MVATQISWPPTIIFRSTDDG	334	

Db 1769 LLASITGHSEPVDFIPGGLIIPAMNPLSINLSCGTGATVPAITIGIPDLVAHSLTG- 1827
Qy 335 ATWRIWMTSYNRSRYLVLDISAEPWLTFCVQNPVPSPKLGWDE- 383
Db 1828 -----PVHILIDLPAVP--GFG-----NTTGAPSSGFFNSGAGGVGFGNV 1866
Qy 384 -AMAIIDPFN-SDRMLYGTGATLYATNLTWKDSCQGHIAIPMWKGLBETAVNDLISPPSG 441
Db 1867 GAVMSGWGNQAPSAALLGGSGVFNAGTL-----HSGVLNFGSGMSGLFNTSLVLG- 1917
Qy 442 AP-LISALGDLGGTHADVTAVPSTIFTPVFTTGTSDVYAEINPSIIVRAGSDPSSQP 500
Db 1918 APALVSLGSGVQ-----QLSGLLASGTA-----LHQGLVLPGLADVGL-- 1957
Qy 501 NDRHVAFTDGGKMWFOGSEPGVTTGGTVAASADGSRFV-WAP-----GD-----PGQPVVY 552
Db 1958 --GNVGLGNVDFNLGAGNVGFGNVGGNNGVGNVGLGNVGMNFGNLGSLTGLMGLG 2015
Qy 553 AVGFNSWAAASQGVPAANAQIRSDRVNPKTFVALSNGTF-----YRSTGGV 598
Db 2016 NIGFGNAGSYNFGI-ANM-----GVGNIGFANTGSGNFGIGLTGDNLTGFGGFTGSGNV 2069
Qy 599 TFQPVAAGLPSSGAVGV-MHFAVPKEDLWLAASSGLYHSTNGSSWSALTGVSSAVNV 657
Db 2070 -----GLFNSGNGVGNFGNFGTGN-----WGVFNAGSY-NTGIGNSGLASTGLFNAGGF 2117
Qy 658 GFGKSPAGSSYPVAVVVTGGTGVAYRSDDCGTTWVLINDQHQYGNWQA-----ITG 712
Db 2118 NTGVNAGSYNTGSFNAGQAN--TGGFNPSGVNTGMLTGDINTGVANSGDVNTGATFISG 2175
Qy 713 DHANLRVYITGNRGIVYGDIG 735
Db 2176 NYSN-----GAFWRGDYQGLLG 2192

RESULT 2

US-11-052-554A-281
; Sequence 281, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR FILING DATE: 2004-07-20
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 281
; LENGTH: 5291
; TYPE: PR1
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-281

Query Match 4.2%; Score 168.5; DB 7; Length 5291;
Best Local Similarity 20.1%; Pred. No. 0.0031;
Matches 190; Conservative 82; Mismatches 270; Indels 401; Gaps 46;
Qy 27 GARGILYVRDTCGMYRWDAANGRWIPLLDWGNWNNYGVVSVIAADPINTKNVWAAVG 86
Db 4186 GANEFAQISTDNGA-----TWNV-----VTVAADSLN----- 4212
Qy 87 MYTNSDPNDGAILRSSDQATWQITPLPKLGNMFGRGERLAVDPNN----- 137
Db 4213 -----WSYVDGRILTNGT--TTWQVRVV--DLAGNV--GATSSOSALIDTVPNAQVLTIAS 4262
Qy 138 -----DNILYF-----GAPSGKGLMRSTDGATWSQMTNFPDVGTYIAN 176

Db 4263 ISTDTGSSATDFITSDTMLTLTGLSILGAGLASEVAQISLDSGATWTLTTNGTQWITYTDS 4322
Qy 177 PTDFTG-----YQSDIQGVVWVAFDSSSSSLGQASKTIFVGVADPN---NPFVWSRDCGATW 230
Db 4323 RLTLDGSVYQVRV-----LDLAGNTGPVVKTKVVDVTINPTATPTIVSYTDDVGQRQ 4375
Qy 231 QAVPGAPTGFPHKGVDFPVNHLVLIATNSNTGPPYDG-----SSGDV----- 272
Db 4376 GTLSSSQ-----ATDDTPLLNGVLASPLASGEVVLYLRNGILLGA 4416
Qy 273 -----WKFS-----VTSQWTW---RISVPSTDDTANDYFGYSGLTIDRQHENTIMVAV- 317
Db 4417 VTMVGLNWTYSDSLGVSQAYTYSARVVDLAGNITSSDF---VLTVDTSIPTTLAQITS 4473
Qy 318 -----QISW-----PDT----- 325
Db 4474 QTRDTPTIISGVITAALASQYVEVINGKTYTSEPQGVVVDPAHNTWYVQLPDTDAL 4533
Qy 326 -----IIFRSTDGATWTRIWDWTSYPNRSRLRYVL 355
Db 4534 TVSATVTVTAQVKSSAGNGNNANISNGTVTVNAADYTPVTWTASKTTAW---GLTYGL 4590
Qy 356 DISAEPWLTFCVQNPVPVPSKLGWDEAMALDPNSDRMLYGTGATLYATNDLTKWDSG 415
Db 4591 D-SHGMVTVLANQOVMQSTDP-LTWSKTALT-----LYQSGNN-YATSSIAIDYDRN 4638
Qy 416 QOIHIAPMWKLEETAVANDLISPPSGAPLIS---ALGDLGFT-HADVAVPSTIFTSP 470
Db 4639 G-----TGDLFITRDDYGTGYINGFTNGDGT-----FSSA 4669
Qy 471 V-FTTGSVDYAEINPSIIVRAGSDPSSQPNDRHVFST--DGKNNWFQSGEPG----- 522
Db 4670 IQVTVGTLTWYGS-----VAFDKEGDGYLDFEWIGDAGGPDST 4708
Qy 523 -----GVTTGCTVAASADGSRFVWAPGDPQVYVAVFGNSWAAASQGVPAANAQIRSDR 576
Db 4709 FLWNAGTLVGNSTNSGGSATVGG-----AVTGLSLNEGSGVDLNDGRIDL 4758
Qy 577 V-----NPKTFVALS-----NGTFVRSVDGVTQFVAAAGLPSSGAGVGMFHAVPK-BG 625
Db 4759 VOHTYNLNYYTLSSLINQNGTTFWQONTNTFLSGAGSGAMSSSVMTWADFDDGDM 4818
Qy 626 DLWLAASSG-----LYHSTNG----- 641
Db 4819 DLELPASQGRANYGSLLFNNGVLCVPVAGTATTYASQFSLAVDWNHDLGLDIARIAQ 4878
Qy 642 -GSSWSALTGVSSAVNVGFGKSPAGSSYPVAVVGTIGGVT-----GAYR---SDDCG 690
Db 4879 TGQSY-LYTNVSNASN--WTQSALGSSQS-----GTTSGVAAMDYWDGAVDVLVSKQSG 4930
Qy 691 TTWVLINDQHQYGNWQAITGDHANLRVYVIGTNGRGIVYGD 733
Db 4931 SVFLSRNTNTVSYG-----TSLHLRI-----TDPNGINYYGN 4963

RESULT 3

US-11-052-554A-17
; Sequence 17, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR FILING DATE: 2004-07-20
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 17

Db 1867 TSLGSLNAGTASGLVNLNGSGTAISGLFNNVSTLDTATPAVISGFSNLGDHMSGV---SIDGL 1923
Qy 598 V---TFOPV-----AAGLPSSGAGV-----MFHAVPGKGD 626
Db 1924 IAILTPPABSVFQIIDAIAELQHLIDGNALALGNVGGVNLGLANVGFNLGAGNVGN 1983
Qy 627 LWLAASGLYHSTNGSGSWAITGVSSAVNVFGKSAFGSSYPAVFVVGTIGGVTGAYRS 686
Db 1984 INVAG-----NLGSLNGLGNVGTG-NLGFNGNAGN-----FGFNGAGLTAGAGGL 2030
Qy 687 DDCGTTWVLINDDOHQVGNWQAITGHNLRVYVIGTNGRGI-VYGD-----IGGAPSG 740
Db 2031 GNVG-----LGNAGS---GSWGLANVG-VGNIGLANTGNTGIGLTGDRYTGIGGLNSG 2081
RESULT 5
US-11-052-554A-134
; Sequence 134, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052.554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 134
; LENGTH: 2204
; TYPE: PRF
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-134
Query Match 3.9%; Score 156.5; DB 7; Length 2204;
Best Local Similarity 21.18; Pred No. 0.0072;
Matches 177; Conservative 91; Mismatches 287; Indels 283; Gaps 42;
Qy 9 SNVAICGGGFVDGIVFNEGAPGILYVRTIDCGMYRWDAAAGRWIPLLDWVGWNNWG-YN- 66
Db 100 NNVGIGNTGTFIDVCLANLGSYNIgf--GNLGD-----DNLGFGNFGSYNI 142
Qy 67 GUVSIAADPIN-TNKWAAVGM-----YTNWDPNDGAILRSSDQGAQTWQITPLPF 116
Db 143 GFGVNGNDNLGFANAGGNGTGFANTGSNNVGFNGTGSNNVIGLTGNGQ----- 191
Qy 117 KLGNNPGRGMGERLAVDPNNNDILYFGAPSGKGLWRSTDGSAWTSQMTNFPDVCVTIAN 176
Db 192 -IGFSGNSGSGNIGLFGNSGNIgfN--SGSNGFIANSGSFNTGIGNTNTGTLFN 248
Qy 177 PTD-TTGYQSDIQGVVWVAFDKSSSLGQAKTIF-VGVADPNPNVFWMSRDRGGATWQAVP 234
Db 249 SGDVTNG-----AFNPGSFNTGSFNTGSFNTGNGFNGN----- 281
Qy 235 GAPTGFIPHKGVDPNVNHLVYATSTNGGYPDGSSGDVWKPVSVTGWTWTRISPVSTDTA 294
Db 282 -TWTGY-----LNIQNYNTG---IANTGDVDTGAFITGNYSN-----GLFLS 319
Qy 295 NDYFGYSGLTIDRQHPNTIMVATQISWWPDTILFRSTDGGATWTRIDWTSYPNRSLRYV 354
Db 320 GDYQGLVGLMLVDMPILSGLVNI-----PIDIPITASAGNITLMGV---TIPPTGDI--V 371
Qy 355 LDISAEPLWTFG--VQNPVPVSP-----KLGWMDMAIDPFNSDRMLYGTGATLYATND 408
Db 372 LSSIAQORAHFGPITIPNITVGTPTTVAIGPNTAITI-----TG----- 412
Qy 409 LTKWDSGGQIHI-----APMWKLEETAVN-----DLIS 437

Db	413	-----GGAIRIPLISIPAAFGNGSNSTPNPSSGFFNTGAGGASGFGNFGGANSQFWNLAS	466
Qy	438	PPSGAPLISALDGLG-----GFTHADVTAVPSTTFTSVPVTTGTSTV-----	478
Db	467	ATSGASGLLNVGALGSLANVGTTSVGFYNTSTDLATPAFNSGLANISTSIAGLLRST	526
Qy	479	-----DYAELNPSIIIVRAGSPFSSQPNDRHVAFTDGGKNW-----FQSEPGGV	524
Db	527	GTWVLNLGLANHGTLVNGI-----ANLGDYNI GFANLGSANFGSANIGNNIGGA	576
Qy	525	TTGCTVAASADGSRFWAPGPGQPVV-----YAVFGNSWAASQGVPAANAQIRSDRV	577
Db	577	NTGIFDGLANLGSYNTGFGNFGDNELGFGNLGSYNTGFGNLGNDNLGF-ANTG-----	629
Qy	578	NPKTFYALSNGTFRYSTDGGVTFQPVAAGLPSSGAVGMFHAVPKEGDMLW-----AASS	633
Db	630	-----SNNI GFANTGS-----NNIGLGTGDOIG--FGSLNSGSGNIGLFGNSGNI	675
Qy	634	GLYHSTNG-----GSSWSAITGV--SSAVNVVGFGKSA-----PGSSYP	669
Db	676	GFNFGNGNVGICNTGTANFGLGNTGTNTGTFNFGSDVNTGIGNTGTSFNTPGDSNT	735
Qy	670	AVF-----VVGITGV-TCAYRSDC--GTTWVLNDDOHQYGNWGOAITGDHANL	717
Db	736	GDFNPGSYNTGLGNTGDTGDTGAFISGSYNGFLW-----SGNY-QGLIGLHAAL	783
RESULT 6			
US-11-052-554A-142			
; Sequence 142, Application US/11052554A			
; Publication No. US2005028866A1			
; GENERAL INFORMATION:			
; APPLICANT: Sachdeva, et al.			
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE			
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL			
; FILE REFERENCE: 30853/40359A			
; CURRENT APPLICATION NUMBER: US/11/052,554A			
; CURRENT FILING DATE: 2005-02-07			
; PRIOR APPLICATION NUMBER: US 60/589,227			
; PRIOR FILING DATE: 2004-07-20			
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004			
; PRIOR FILING DATE: 2004-02-06			
; NUMBER OF SEQ ID NOS: 763			
; SOFTWARE: PatentIn version 3.3			
; SEQ ID NO 142			
; LENGTH: 3157			
; TYPE: PRT			
; ORGANISM: Mycobacterium tuberculosis H37Rv			
US-11-052-554A-142			
Query Match 3.9%; Score 156; DB 7; Length 3157;			
Best Local Similarity 20.3%; Pred. No. 0.012;			
Matches 174; Conservative 96; Mismatches 295; Indels 294; Gaps 42;			
Qy	13	IGGG-----GFVDGIVFNEGAPGILYVRTDGGMYRWDAANGRWITPLLDWVGNWNGYNGV	68
Db	811	VGGNGVFGNIGANDFGLGNAGLAAGLAGVGNICLGNAGSN-----VGFGNWGVGNI	863
Qy	69	VSTAADPINTKNVAAVGMVNTSNWDPNDGAILRSSDQAGATQITPLPFKLGKGNMPGRG-M	127
Db	864	-----GFEGTGNLGLTCDNQTG-----IGGLNSGAGNI	895
Qy	128	GERLAVDPNDNIIYFCAPSGK-GLWFS-----TDSGATMSQ-WTNFPDVGVTYIANP--	177
Db	896	G---LFSNGTGTVGLFNSGTFNFGNFGSGFTNGIGNGGTGSTGLFNAGNFTGVANPGS	952
Qy	178	-----TDTTGYSQSDIQGVVWVAFDKSSSLGQASK-----	207
Db	953	YNTGSEFNVGDTNTGTFNPGSINTCW--FNTGNANTGVANSNVDTGALMSGNFSNGILWR	1010
Qy	208	-----TIFVGVADPNPNVFWVRDGGATWQAVPGAPTGFIPHKGVFDPVNVHVLVYATSN	260
Db	1011	GNPEGLFGLNVGITIPEFPFTHWTSTGII-----GPI-IIPDITLIPPI-HLGLTGOAN	1061

QY 261 TGGPYDGGSSGDKWKFSTSGTWTRISPVSTDDTANDYFGYSLGTIDRQHPNTIMVATQIS 320
Db 1062 YG-----FAVDP-----IPAI-----HIDFG-----1080
QY 321 WMPDTIIPRSTGGATWTRIMDWTSPYNSRSLRYLVIDISAE-----PWLTFGQPNPPVSP 376
Db 1081 -----ADAGFT-----APATLLSALGTGQFRPGPITVSNVQLNPPFNVN 1122
QY 377 KLGWMEAMA---IDPFNSDRMLYG---TGATLYATNDLTAKWDSGGQIHIAPMVKGLEET 430
Db 1123 KLOFLHDAFFNEFPDPTISVQIQAIPLTSATL-----GGLALPLQOTIDAELP 1172
QY 431 AVNDLISPPSAPLIS-ALGDLGGTHADVTAVPSTIFTSPVFTTGTSDVYAEALNPSIIV 489
Db 1173 AIFSQSIPIDIPIDIPASTINGISMSEVPIDVSDIPAVTITGTTRIDRPLNFDVLS 1232
QY 490 RAGSPD-----PSSQPNDRHVAFSTGGKNWFGSGEPGVTGGTVA 531
Db 1233 SAGPINISIIDIPALPGFNSTELPSG-----FFNTGGG-----GGSGI 1272
QY 532 ASADGSRFVWAPGPGQPVVYAV-CFGNSWAASQGVANAIQIRSDRVNPKT-----581
Db 1273 ANF-GAGVGLLNQASSPMVGTLSGLNAGSLASGLVNSGVDISGMFNVSTLGSAPAVIS 1331
QY 582 -FYALSN-----GTYRSTDDGVTQPVAAGLPS--SGAVGMFHAVP-----621
Db 1332 CFGNLGNHVSQVSDGLLAMLTSQ-----SGSGQPSIIDAAIAELRLNPLNVLN 1387
QY 622 -----GKEGDLMASSGLYHSTNGSSWSAITGVSSA--VNVFGKSAFG-----S 666
Db 1388 GSYNLGFANVGDVNLGAGNLGNLGNLGNLGNLGNLGNLGNLGNLGNLGNLGNLGNLGN 1447
QY 667 SYPAVVGVTIGGVTGAYRSDCGTTWVLINDDQHQYGNWQAITGDHANLRVYVIGTNG 726
Db 1448 GLGALPGIGNIG--LGNAGSNVVG-----FGNMGU-----GNIGFGNTGNN 1487
QY 727 RGI-VYGD-----IGGAPSG 740
Db 1488 LGIGLTGDNQTFGGLNSG 1506

RESULT 7

US-11-087-099-7521
; Sequence 7521, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 7521
; LENGTH: 7465
; TYPE: PRT
; ORGANISM: Magnetococcus sp. MC-1
US-11-087-099-7521

Query Match 3.8%; Score 155; DB 7; Length 7465;
Best Local Similarity 19.7%; Pred. No. 0.041;
Matches 153; Conservative 69; Mismatches 287; Indels 268; Gaps 31;

QY 76 INTNKKV-----AAGVMTNSWDPNNDGAILRSSDOGA---TWQITPLPFKLGNNMPG 124
Db 6509 VGSNGVMSYQLDNSKAVVQGLTQGOQVSESFTVHTSDGSSQLTNIT-----6556
QY 125 RGMGERLAVDNNNIIYFGAPSGKGLWRSTDSGATWSQMTNFPDVGTYIANPTDTTC-- 182
Db 6557 -GRQDSAVGSGTGVSTEDSTLSSSGMLSVSDADSGEAAVYASTQTSYGSFSLDSSGAW 6615
QY 183 -YQ-----SDIQGVVWVAFDKSSSLG-QASKTIFVGVADPN-NPVFWSRDGGATWQAVP 234

Db 6616 SYQLNNSASQVQALI-----MGQVSSBSFTVATVDGTASSVVVTVVGAQDAAQLG 6665
QY 235 GAPTGFIPHKGVDFDNNHV-----LYIATSNVG--GPYDSSGDDWKFVSFTSGTW 282
Db 6666 GYQSGVGVEDGQLSTAGQLTISDADAGQLYQVASTDGMYGHPAMDSSGAWSY-VLNNE 6724
QY 283 TRISVPSTDDTANDYFGYS-----GLTIDRQHPNTIM---VATQISWWDPTIIFRS 330
Db 6725 AAIQMIAGTQVLESFTVSSSDGTQAVSVTITGSDNAVITGEAAAGVSETNDDSSMA 6784
QY 331 TD-----GGATWTRIMDWTSPYNSRSLRYLVIDISAEPLWTFGQPNPPVSP 375
Db 6785 TCKLINSIDLSQAHFVGPGQAGVCAFTLOGDGNWTVLD-----NHGV-----6829
QY 376 PKLGMWDEAMAIIDPFNSDRMLYGTGATLYATNDLTAKWDSGGQIHIAPMVKGLEE- 429
Db 6830 -----AMA-----GLGTGDVVTFVRSILDNTAETVITTINGLDEPPTAEE 6872
QY 430 -----TAVNDLIS-----PPSGAPLISALGDL-----451
Db 6873 IEARLABEQRLOQLQASNDLLOGAIGAEGGNAGGTAPPPNGNPNADGAPDGLGAGGA 6932
QY 452 -----GFTHADVTAVPSTIFTSPVFTTGTSDVYAEALNPSIIVRAGSDPSSQPNDRHV 505
Db 6933 AAGGATGGTGGTTA-----GTGGGGLG-----GGFG-SCAPNT---6967
QY 506 APTDGGKNWFOGSEPGVTTGCTVAASADGSRFVWAPGDPQPVVYAVFGFNSWAASQ 565
Db 6968 --SAGLTTGGGTATGTTTGT-----GGTGGNATG 6998
QY 566 VPANAQIRSDRVNPKTFYALSNGTTFYRST--DGVTFQPVAAGLPSSGAVGMFHAVPGK 623
Db 6999 GTA-----TGTATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 7034
QY 624 EGDMLAASSGLYHSTNGSSWSAITGVSSAVNVNFGKSAAPSSYPVFTVGTIGVTGA 683
Db 7035 AGGTGGAAGGAATGEAGGTATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 7089
QY 684 YRSDCGTTWVLINDDQHQYGNWQAITGDHANLRVYVIGTNGRGIYVDIGGAPSG 740
Db 7090 AATGEAGGT-----ATGTTGTTATGGTA-----TGTGGAAGTGGAAAG 7129

RESULT 8

US-11-087-099-12161
; Sequence 12161, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 12161
; LENGTH: 807
; TYPE: PRT
; ORGANISM: Cytophaga hutchinsonii
US-11-087-099-12161

Query Match 3.8%; Score 153.5; DB 7; Length 807;
Best Local Similarity 17.9%; Pred. No. 0.0034;
Matches 160; Conservative 108; Mismatches 223; Indels 403; Gaps 52;

QY 8 WSNVAIGGGGFGVDGIVFNEGAGIILYVRTIDIGMYRWDAANGRWIPLLDWVGWNNW--- 63
Db 31 WKWA-----KGSHTTAVRQD-CTLNSW-----GQMYGQLG 62
QY 64 ---GNGVVSAAADPINT-NKMWAAVGMVTNSWDPNNDGAILRSSDOGATW----- 109
Db 63 DGSYSSIIIPVQVGLTWTWVKVFA---KYDNSF-----AIMK---DGTLWAGMNSLIGTL 111
QY 110 -----QITPLPFKLGNNPGRGMRERLAVDPNNNDILYFGAPSGKGLWRSTDSGATWSOM 164

Db 112 GDGTEQVRKTVKVG-----TATD-----WQSV----- 135
 Qy 165 TNFPDVGTYIANPTDTTGYQSDIOGVVWVAFKSSSLGQASKTIIFGVADPNPVPFWSR 224
 Db 136 TNF-----NFTAIAKKD--GSLWIDGNYNGQNGTKV-----KNLIPSKV 175
 Qy 225 DGGATWQAVP-GAPT-----GPIPHKGVDPV----- 250
 Db 176 GSGKNWKSVSIGSSNLAIAKTDGLWGMNYYNQIGYPTPTVDLSPTQIGSETDWSVS 235
 Qy 251 ---NHVLVIATSN---GGPYDGSSGD-----VWKFVS 277
 Db 236 CGDHHVAIRTNGSLWWSKEQIGNEINSYQTEPVQVGSDDTNGTCHAIQDWSIAI 295
 Qy 278 -TSCT-WTRISPVSTDTANDFYGSGLTIDRQHPNTIMVATOI-----SW---WP---DT 325
 Db 296 KTGSLWTWGS-----YGYNSSYIRKN-----IPTQGRDTSWMIYPSQOS 340
 Qy 326 IIPRSTGGATWRIWDWS-----YPNRSRLRYLD-----ISABPW-----L 363
 Db 341 VAFK-TDGS-----LWTGWSNARQLGLGLYCNINIPYVWSCPEVLTSSCWQAASIYQST 394
 Qy 364 TFGVQPNPPVPSKLGWDEAMATDPNSD-----RMLYGTGATLYATNDLTK 411
 Db 395 SFGURTDSTL--WKWGWGNE-NSLDIYARNPDGPFARIGTYKRAYASGGALLAKDDGTL 451
 Qy 412 WDSGGQIHIAPMVKLEETAVNDLISPPSGAPLISALGDLGGFTHADVAVPSTIFTSPV 471
 Db 452 WQSGR-----NVPVL----- 462
 Qy 472 FTTGTSVDYAEIAPNPSIIIVRAGSFPDPNDRHVAFSTDGG-----KNMFQSGEPG-- 523
 Db 463 -LSSSKDWKAAVAS-----SSRGVAIKTDGSLWGDIDNYNNTEPSSGV 506
 Qy 524 -----VTGTGTAASA--DGRFWAFDGPQPVVYAVFGNSWAAASQGV 566
 Db 507 IPLLKEINPGTQWQSISASTTTAAAIARDGSLWIW-----GSALYGAMGTG---IAVAGS 558
 Qy 567 PANAOIRSDRVPKTFYALSNGTFYRSTGGVTFQPVAAAGLPSSGAVGVMFHVPKGEKD 626
 Db 559 PTLIQMND-----TGQFVSGESTTMAIKT-----DGT 588
 Qy 627 LWLAASGLYHSTNGSSWAITGVSAVNVVFG---KSPAGSSYP---AVFVGTIGGVT 681
 Db 589 LWACGQNNYQGLNGNT-----TDIYTLTQIGTATDWTWVAGPYHTLAIKTDSING-W 642
 Qy 682 GAYRSDCGTTWVLINDDQYQWNGQAITGDHANLRVYGTNGRGIVY--GD 733
 Db 643 GSYTYNKLGLGLDYGLNLEQ--NWKHIATGPYNS-----FAIHGDGTLTYTVD 689

RESULT 9

US-11-087-099-11214
 ; Sequence 11214, Application US/11087099
 ; Publication No. US20060041961A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Abad, Mark S. et al.
 ; TITLE OF INVENTION: Genes and Uses for Plant Improvement
 ; FILE REFERENCE: 38-21(53450)B EP
 ; CURRENT APPLICATION NUMBER: US/11/087,099
 ; CURRENT FILING DATE: 2005-03-22
 ; NUMBER OF SEQ ID NOS: 12464
 ; SEQ ID NO 11214
 ; LENGTH: 894
 ; TYPE: PRT
 ; ORGANISM: Cytophaga hutchinsonii
 US-11-087-099-11214

Query Match 3.7%; Score 148.5; DB 7; Length 894;
 Best Local Similarity 20.0%; Pred. No. 0.0086;
 Matches 145; Conservative 87; Mismatches 238; Indels 255; Gaps 39;

Qy 7 TWSNVAI-----GGGFVDGIVP-----NEGAPGI-----LYVRTDIGMYRDAAN 48
 Db 203 TWTGIAACFSSGIGIKANGTLTWGPNNGGELGIGNDEKHTFPVQVGADQNMLYAVAGSN 262
 Qy 49 GRWIPLLDVGWNNWNGYGVSI-----AADPINTKVAAGVMTN----- 90
 Db 263 HIVALKVDGSLW-GWGLNEAGEIGNGTVQODSPVRVGTDDNDWTTLAAGSNHTLAIKSNG 321
 Qy 91 ---SWDPNDGAILRSSDQGAWQIITP---LPFKLGNMP-----GRGGERLADVDPNND-- 138
 Db 322 TLWAGHN---VRGNLGNATTLPATLPVQVGTADWIRVSAGLDHSAVALKANGTLWA 377
 Qy 139 ---NI---LYGAPSGKGLWRSTDSGATWSQMTNPDVGTIYIANPTDTTGYQSDIOGVVW 192
 Db 378 WGVNVRGOLGYGSGTDKHLVPQIGTAHTWVAI-----NAGTY-----HTIGVKAD--GSLW 426
 Qy 193 VAFKSSSSLSGOASKTIIFGVADPNPVPFWSRDCGATWQAVP---GAPTGTIPHKGVPD 248
 Db 427 TWGNTNAQLG-----DGGTADQPVPHSIRTAPDDWIS----- 459
 Qy 249 PVNHVLIATSNITGGPYDGSSGDVWKFVS---TSCTWTRISPVSPSTDTANDFYGSGLT 304
 Db 460 -----IAASNAFSMGLKANGTINTWGLNPFETDGOYKNSSPV---QTGSD----- 501
 Qy 305 IDROHPNTIMVATOISWWPDTIIFRSTGGATWTRIMDWTSYPNRSRLRYVLDISAEPWLT 364
 Db 502 -----QNWKSIATGSNY---ILALKADG---TLWAGD-----N 529
 Qy 365 FGVQPNPVPSPKL-----GHWDEAMALDPNSDRMLYGTGATLYATNDLTKW--D 413
 Db 530 FSGQLGDCGTEQPRMLPKOIGTATTWVSTAAA---SGVQSFG---IQADGSLWSWGHN 580
 Qy 414 SGGQIHIAPMVKLEETAV---NDLISPPSGAPL---ISALGDLGGFTHADVAVPSTIF 467
 Db 581 YDGLGLGTNLKTLVPRFVGTGTWISIKTVGHTLGIKSDGTLWGW----- 627
 Qy 468 TSPVFTTGTSDYVLAELNPSIIIVRAGSFPDPNDRHVAFSTDGKKNWFQSGEPGVTTG 527
 Db 628 -----GTN-----RQQLGDGTLVN-KYVPVQLGTDTRDWIN-----IAG 660
 Qy 528 GT---VAASADGRFVWAPGDPGQPVVYAVFGNSWAAASQGVPANAOIRSDRVPKTFYA 584
 Db 661 GTIHNICLKSNTLWAGRNNGYQ---VGTGNLITQDRMVQISAE-----QNNIDVYA 710
 Qy 585 LSNCTFYRSTGGVTFQPVAAAGLPSSGAVGVMFHVPKGEKD-----LWLAASS 633
 Db 711 SLDSQFAIRADGSLW---ACGLNSSGQLG-----DGTSDVRAAMVAVGTNQKWISLAA 760
 Qy 634 GLYHS 638
 Db 761 GNYHT 765

RESULT 10

US-10-517-939-330
 ; Sequence 330, Application US/10517939
 ; Publication No. US20060003433A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Steer, Brian
 ; APPLICANT: Callen, Walter
 ; APPLICANT: Healey, Shaun
 ; APPLICANT: Hazlewood, Geoff
 ; APPLICANT: Wu, Di
 ; APPLICANT: Blum, David
 ; APPLICANT: Esteghalalian, Alireza
 ; TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM
 ; TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
 ; FILE REFERENCE: 56462007901
 ; CURRENT APPLICATION NUMBER: US/10/517,939
 ; CURRENT FILING DATE: 2004-12-13
 ; PRIOR APPLICATION NUMBER: PCT/US03/19153
 ; PRIOR FILING DATE: 2003-06-16
 ; PRIOR APPLICATION NUMBER: 60/389,299

; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 380
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 330
; LENGTH: 755
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample.
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(35)
US-10-517-939-330

Query Match 3.7%; Score 147.5; DB 6; Length 755;
Best Local Similarity 20.4%; Pred. No. 0.0082;
Matches 159; Conservative 74; Mismatches 254; Indels 291; Gaps 41;
QY 44 WDAANGRIPIPLDVGWNNWNGVGVVSTAAADPINTKVAAGVMTNS---WDPNDGAIL 100
DB 123 WHSQNPGLTN-----GNWSRSQLINIMNDHIDT-----VAGRYAGEVLVWD---VYN 167
QY 101 RSSDQATWOTLPLFKLGGNMPGRGMRERLAVDP-----NNDNLYFGAPS----- 147
DB 168 QAFNEDGTYRSTIYNGIGQYIDIAFTRARAADPHAKLIYNDYINIGLMSKSNVYNA 227
QY 148 -----GKG-----LWRSTDSGATW-SQNTNFPDVG--TYIA-----NPTDTT 181
DB 228 ADMVRGVPIDGVGFQMHLEGGVSGSSLASNMORFADLGLVYITELDVRIQNPNT--- 284
QY 182 GYQSDIQGVVWVAFDKSSSLGQ-ASKTIFVGVADPNPNFVMSRDRGGATWQAVP----- 234
DB 285 --QODLQAAAYQVTVNRCLAQAPACKALQV-----WGIPDKYSW--VPDVFPGT 330
QY 235 GAPTGFIFH---KGVPDPVNHVLYIA-----TSNTGPPVGGSGDVWKFVSIGTWTRISP 287
DB 331 GAPLLFNNDYBAKPAYAVQAEQELMAANPQNTPGTPAHTPS-----ATSTSAATATPP 384
QY 288 VPSDTANDYFGYSLGTLIDRQHPNTIMVATQISWPDITIIIPRSTGGATWIRINDWTSYP 347
DB 385 ATATATATTPSGGGVCAVD----- 403
QY 348 NRSRLYVLIDISAEPLWLTGQVQPN-----PPVSPKLGWMDAMADPNFSDRMLYGT 399
DB 404 -----YVI---ANQWGN-GFOANVTIITNHSAPVNGYTLAWTHAPQOIVTSQWNTIAQS 454
QY 400 GATLYATNDLTKWD-----SGGQIH-----IAPMVKGLBEETAVNDLISPPSGA 442
DB 455 GSASASNPAGYNGVIGANGGKISFGQGLAGGSAAVAPTYFALNGAACNGAVLPPTAT 514
QY 443 PLISALGDLGGFTHADVT-----AVPSTIFTSPVFTTGTSDVYAEALNPSIIVRAGSPDPS 497
DB 515 -----FTSPPTATMCPQATPELLVQPVTSPTT-----QLSQTLVRLNGNS-- 556
QY 498 SOPNDRHVAFTDGGKMFQSGEPGGVTTGGTVAASADGSRFVWAPGDPQGVVYAVGFG 557
DB 557 -----WRAAGPAGVVT-----VTAP-DPD-----GYF 578
QY 558 NSWAASQGVPAQAQIRDRV-----NPKTP--YALNSGTIFYRSTDDGGVTFQPVAAGLPS 609
DB 579 RLTIPLAANTSNAILVEGRVVRVITHSNGCTYGGYTLRSRTVTVIQVASSPVLTPATPSPT 638
QY 610 -----SCAVGVNF-----HAPVQKE--G 625
DB 639 ATATPTVTATPSGACTVAYAITNDWGSFGFTANVTLTNTGGSALMGWTLAYAFQGNQTI 698
QY 626 DLW--LAASSGLYHS-TNGGSSWSAITGVSSAVNVGFKGSAPG--SSYPAVFVVGITGG 679
DB 699 NAWNGTAVQSGSSVSVTNAG--WNGSLPPNVASFGQASYSGNNSVFPASF---TLNG 751

RESULT 11

US-11-052-554A-6

; Sequence 6, Application US/11052554A
; Publication No. US20050288666A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 949
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-6

Query Match 3.6%; Score 146; DB 7; Length 949;
Best Local Similarity 20.0%; Pred. No. 0.014;
Matches 147; Conservative 77; Mismatches 252; Indels 258; Gaps 37;
QY 136 NNDNLYFGAPSG---KGLWRSTDS---GATWSQMTNFPDVGTYIANPTDIT--GYQSD 186
DB 70 NHDNQIVCTANGMTITSTGLGLPDSSEENTGQMIQN-----GGIAGNTVTTNGHQVV 123
QY 187 IQGVVWVAFDKSSSLGQASKTIFVGVADPNPNFVMSRDRGA-----T 229
DB 124 LEG-----GTASDTVI-----RDGGQSLNGLAVNTTLNRRGEQ 157
QY 230 WQAVPCAPTGFPHKGVPDPVNHVLYIATNTGPPVGGSGDVWKFVSIGTWTRISPVP 289
DB 158 WHEGCVATGTINRDGYQSV-----KSGG-----LATGTIINTCAEG 195
QY 290 STDITANDYFGY-----SGLTIDRQHPNTIMVATQISWPDITIF---RSTGGATWTR 339
DB 196 GPDSDNSYTGQKVGQTAESTTINKNGRQIILFSGLAR---DTLIYAGGDSVHGRALNTT 252
QY 340 IWDTSYNRSLRYVLIDISAEPLWLTGQVQPNFVSPKLGWM----- 381
DB 253 LGGYQYVHRD-----GLALNTVINEG--GWQVVKAGGAAGNTTINQNGE 295
QY 382 -----DEMAIDPFNSDRMLYGTGATLYATNDLTKW-----DSGQIHIAPM 423
DB 296 LRHVAGGEATVQNTGGALVTSTAATVIGTNRNLGNFTVENGGKADGVVLESQGRLDV--- 352
QY 424 VKGLEETA VNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGT-----S 477
DB 353 ---LSSHSAQNTLVDDGGTLAVSAGKA---TSVITSGALLIADSGATVEGNTASQKFS 406
QY 478 VD--YAEALNPSIIVRAGSPDPSQPNDRHVAFTDGGKMFQSGEPGGVTTG--GTVAAS 533
DB 407 IDGTSQASGLLENGGSGFTVNA-----GGQAGNTTVHGRGTLTLA 447
QY 534 ADGSRFVWAPGDPQGVVYAVGFGNSWAASQGVPAANA-QIRSD-RVNPKTFFYALSNCTFY 591
DB 448 AGGSLSGRTQLSKGASMWL-----NGDVVSTGDIVNAGEIRFDNQITPANA--ALSRAV-- 498
QY 592 RSTDGCVTFQPV-AAGLPSSGAVGVWFHAPVKGEGDLWLAASGLYHSTNGSSWSAITG 650
DB 499 AKSNSPVTPHKLTNTLTQGGGTINMRVLDGNSASDQLVINGG--QAT--GKTLWLAFTN 554
QY 651 VSSAVNVGFKGSAPGSSYPVAVFVVGITGGVTT-----GAYR---SDDCCTTW 693
DB 555 VGNS-NLGVATTGQG-----IRVVDQNGATTGEGAFALSRPQAGAFNTLNRSDSDW 608
QY 694 -----VLINDDHOQYGNWGOAITGDHANLRRYVIGTNG 726
DB 609 YLRSENAVYAEVPLYTSMLTQAMDYDRILLAGSRSHQTG-----VNGENNSVR---LSIQG 660

[illegible]

Db 361 WSRANTKDYNSKILTNQTNLTSTFYTASIGHDYSTGVETRETQTNVG-VNPVTLPAVN 419
Qy 493 SFDSSOPDRHVAFTDGGKQNFQSGSEPGVTTGGTVAASADGSRFVWAPDGPQPVVY 552
Db 420 IYHPDS-----SIHPGGLTRNG--ANANGTDTFA-----IY 449
Qy 553 AVFGNSWAASQGVPAANAQIRSDRVNPKTFYALSNGTFYRSTDDGGVTFQPVAAGLPSSGA 612
Db 450 AF---DTLOITRDPFELNGGIRLDNYHTEYDSATACG---GSGRGAIT---CPAGVAKGSP 500
Qy 613 VGVNFHAPVKGEGDL---WLAASSGLYHSTNGSGSWSAITGVSSAVNVVGFSGKAPGSSVPA 670
Db 501 VTTVDTA---XSGNLVNWKAGA---LYHLTENGNY-----INAVSQQPEGNN-- 544
Qy 671 VFVVGTTGGVGTAYRSD-----DCGTTWVLI-----NDD- 699
Db 545 -FALAQSGSGNSANRTDFKQKANTSEIGTKWQVLDKRLLLTAALFRDIEVEVEQNDG 603
Qy 700 -OHQYG 704
Db 604 TYSQYG 609

RESULT 14
US-11-052-554A-179
; Sequence 179, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 179
; LENGTH: 955
; TYPE: PRT
; ORGANISM: Shigella flexneri 2a str. 2457T
US-11-052-554A-179

Query Match 3.5%; Score 140.5; DB 7; Length 955;
Best Local Similarity 20.1%; Pred. No. 0.033;
Matches 162; Conservative 98; Mismatches 309; Indels 237; Gaps 38;
Qy 77 NTNKVMAVGMVNTSWDPNDGAILR-----SSDQATWQITPLPKLGG- 120
Db 178 NSNGIGTINDYGTSDVLSGSKIKTQSTGVYIGLNGNNAARFTATDLTIDVOG 237
Qy 121 NMPGRGGERLAVDPNDNDILYFCAPSGKGLWR--STDGATWQMTNFPDVGVIANPT 178
Db 238 SAMGINVQKNSVDLTGNTSIKTNGDNAHGLWSFGQVSANALTVDTGAAANGVEVRGT 297
Qy 179 DTTGYQSDIQGVVWVAFDKSSSLGQASKTI-FYGVADPNPNPFWSRDGGATWQAVFGAP 237
Db 298 TTIGADSHISSAQGGGLVTSSS-----DATINFSGTAAQRNSIF-----SGGSYGASACTA 348
Qy 238 TGFPHKGVDPVNHVLIAT---SNTGPGYDGGSGDVWKFESVTSGTWTRISVPSTDTA 294
Db 349 TAVINQMONTITVDNRGSLALGLWALSGGRITGDS-----LAITCAAGAR-GIYAMTNSQ 402
Qy 295 NDYFGYSLRITDRCHPTINWATOISWPPDIIIFRSTDDGATWTRINDWTSYVNRSLRYV 354
Db 403 IDL---TSDVLVDMSTPQMAIATO-----HDDGYAASRI-----NASGRL 441
Qy 355 LD---ISAEPLWTFGVQPNPVPSPKLI-----GWMDEAM----- 385
Db 442 INGSVLKGGILNDMHPGSVWTGSSLSNDVNGKLDVAMNNSVWNTVNSNLDTLALSH 501

Qy 386 -AIDPFNSDRMLYGTGATLYATN-----DL--TKWDSGGQI 418
Db 502 STVD--FASHGSGTAGTFTTLNVLNLSGNSTFIMRADVVGEGVNNRGDILLNLSGSSAGNH 560
Qy 419 HIAPVVKGLEETAVND--LISPPSGAPLISALG--DLGGFTH----- 456
Db 561 VLAINQSEATITGNEVLTVVKTDTGAAASFSSAQVELGGYLDVRKNTNNWELYASGTV 620
Qy 457 -----ADTVAPSTIFTSPV-----FTTGTSDVYAEALNPSII 488
Db 621 PEPTNPPTPAPAPQPIVNPDPTEPAPTPTTADAGNYLVNGLVNYVE-NRTL 679
Qy 489 VRAGSFDSSPNDRHVAFSTGGK--NWFQSGEPG-----GVTTGG----- 528
Db 680 QRMG--DLRNQSKDGNILWRSYGGSLDSFASGKLSGDFMGYSGIQFGDKRLSDVMPLYV 737
Qy 529 -----TVAASADGSRFVWAPDGPQPVVAVVGFSGNSWAASQGVPAANAQIRSDRVNPKTFY 593
Db 738 GLYIDSTHASPDYS-----GGDGTARSDYMGMYASYNAQNGFYSDLVIKASR-QKNSFH 790
Qy 584 ALSNGTFYRSTDDGGVTFQPVAAGLPSSGAVGVNFHAPV-----GKGGDLW 628
Db 791 VLD-----SONGVNANGTANGMSISLEAGORFNLSPTCYGYIEPQTLTYSHQENWA 844
Qy 629 LAASSGLYHSTNGSGSWSAITGVSSAVNVVGFSGKAPGSSYPAVVFVGTIGGTGAYRSD 688
Db 845 MKASNLGLNHLN---HYESLLGRASMI-LGYDITA-GNSQLNVV-----KTGAIR-EF 892
Qy 689 CGTTWVLNDDOHQY---GN-WQQA 710
Db 893 SGDTEYLLNDSREKYSFKGNGWNGV 918

RESULT 15
US-11-052-554A-283
; Sequence 283, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 283
; LENGTH: 1461
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-283

Query Match 3.5%; Score 140.5; DB 7; Length 1461;
Best Local Similarity 18.4%; Pred. No. 0.056;
Matches 148; Conservative 107; Mismatches 304; Indels 245; Gaps 37;
Qy 15 GGGFVDGIVFNEGA-----PGILYVRT-DIGGMYRWDANGRWIPLLDWGVNNGVNG 67
Db 515 GTHVLVTGLVADGSGNNVALDPAV--TRTLDRG-----ANTIFVTVDAAAGTGAASRA 565
Qy 68 VVSTAADPINTNKVAAVGMVNTSWDPNDGATLRSDDGATWQITPLPKLGNMPPGRGM 127
Db 566 ITLGVGSPLIT-----INTVSGDD--IISAEKGA-----PLTLTGSTQOAE 606
Qy 128 GERLAVDPNNNDILYFGAPSGKGLWRSTDSGATWS-----QMTNFPDVGTYI-ANPTDT 180
Db 607 GQTVTV-----TLAGQSFTTTVQADSGSWSLTVPAAAMGNLPDCAVAITASVTDL 655

```
QY 181 TGYQSDIQVWVAFDKSSSLGQASKTIIFGVADPNN--PVFWSRD-----GG 227
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
656 SGNGTNTRTITVDSQAPALSIDPLTADNTINAAESQDLPTGTDAQPGQVTVTLNG 715
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
228 ATQOAVPGAPTGFPHKGVDPNVHVLXIATSNLTGGPYDGGSDVWKFSVTSGTWTRISP 287
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
716 QTYQ-----GVQPGTWS-----VTVPAANVAGALADGNATVTSVNDVAGNPSSVR 763
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
288 VPSTDTANDYFGSLTIDR-----QHPNTIMVATQISWMPDTIIFRSTDGGATWTRIMD 342
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
764 VALVDATPPVVVTINPVATDNTVINPTEHAQAQIISGTVTGAQAGDIVTVTLNNVDYTTVD 823
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
343 -----WT-----SY-----NRSRYLVDISAEPLWTFGVQPNP 371
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
824 GSGNWSLGVPAVVVSGIADGSPVSVSDTKAGNTGQSLLTVVTATA----- 871
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
372 PVFSPKLGWM-----DEAMATDPNSDRMLYGTG-----ATLYATNDLTKWDSGQI 418
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
872 -----PLIGINSIAGDDVINASEKADLQITGTSQDPVNTAITVTLNGQNTYTTTIDASG-- 925
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
419 HIAPMKGLEETAANDLISPPGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSV 478
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
926 -----NWSVTYPA-----SAVTALQOANYTVTAATVSDIGNSATASHNVLV 966
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
479 DYAEIAPSIIVRAGSPDPSPQPNDRHVAFSTDGKKNWFQSGEPGGVTTGGTVAAADGSR 538
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
967 DSA--LPGVTINPVATDDIINAAEAGVAQTISGQVTCAGDGTVTITLGGNTVATVGSN 1024
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
539 FVWAPGDPQPVVYVCGFNSWAASQGVPAQAQIRSDRVNPKTFYALNGTFYRSTDCGV 598
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1025 LTWSVDVPAAD-IOALNGD-----LTYNASVTN-----ONG-----NTGSGT 1061
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
599 TFQPVAAGLP-----SSGAVGVMPHVPKGEGLWLAASSGLYHST----- 639
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1062 RDITIDANLPLGLAVDTVAGDDVNIIEHG-----QALVVTGSSGLAESTPLVTINNVEY 1117
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
640 -----NGSSWSAITGVSSA-----VNVGF-GKSAPGSSY-----PAVFVV 674
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1118 TTAFAQDGSWS--VGVTAQVSAWPACTVNIASVSGESSAGNSVSIHPVTVDLTPAAITI 1175
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
675 GTIGG---VTGAYRSD--CGTT 692
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1176 NTIATDDVINAEEKGADLTLSGTT 1199
```

Search completed: March 2, 2006, 14:28:13
Job time : 20.0053 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006, Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 2, 2006, 12:06:58 ; Search time 266.937 Seconds
(without alignments)
4927.737 Million cell updates/sec

Title: US-09-917-376-3

Perfect score: 4036

Sequence: 1 ATTQPTWSNVAIGGGGFVD.....YGTNGRGIVYDIGGAPSG 740

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/abs/ABSSWEB.spool/US09917376/runat_02032006_091454_8101/app_query.fasta_1
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs806h -USER=US09917376 @CGN_1_1_278 @runat_02032006_091454_8101
-NCPU=6 -ICPU=3 -NO_WMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_THREADS=1 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/1/COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5/COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A/COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B/COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H/COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PCUTUS_COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	392.5	9.7	1103	3 US-09-533-559-7511	Sequence 7511, Ap
c	227.5	5.6	3129	3 US-09-252-991A-13873	Sequence 13873, A
3	227.5	5.6	8211	3 US-09-252-991A-13656	Sequence 13656, A
c	203	5.0	4188	3 US-09-252-991A-13774	Sequence 13774, A
c	201.5	5.0	4131	3 US-09-252-991A-13773	Sequence 13773, A
c	192	4.8	4411529	3 US-09-103-840A-1	Sequence 1, Appli
c	184	4.6	10317	3 US-09-902-540-1027	Sequence 1027, Ap
c	184	4.6	4403765	3 US-09-103-840A-2	Sequence 2, Appli
9	181.5	4.5	11679	3 US-09-328-352-1377	Sequence 1377, Ap

10	177	4.4	29927	3 US-09-949-016-11814	Sequence 11814, A
11	177	4.4	29927	3 US-09-949-016-11814	Sequence 17474, A
12	177	4.4	29927	3 US-09-949-016-11814	Sequence 17475, A
13	177	4.4	37802	3 US-09-949-016-11814	Sequence 12639, A
14	175.5	4.3	4287	3 US-09-902-540-2590	Sequence 2590, Ap
15	175.5	4.3	16584	3 US-09-902-540-1119	Sequence 1119, Ap
16	170.5	4.2	25165	3 US-09-453-702B-39	Sequence 39, Appl
17	170.5	4.2	25165	3 US-10-114-170-39	Sequence 39, Appl
18	170	4.2	4403765	3 US-09-103-840A-2	Sequence 2, Appli
19	169	4.2	4411529	3 US-09-103-840A-1	Sequence 1, Appli
20	168.5	4.2	20113	3 US-09-902-540-1173	Sequence 1173, Ap
21	167.5	4.2	4236	3 US-09-902-540-5367	Sequence 5367, Ap
c	167.5	4.2	34662	3 US-09-902-540-1261	Sequence 1261, Ap
c	164	4.1	2319	3 US-09-252-991A-13875	Sequence 13875, A
24	163	4.0	5741	2 US-07-706-699-4	Sequence 4, Appli
25	163	4.0	5741	2 US-07-998-931-4	Sequence 1, Appli
26	160.5	4.0	47981	3 US-09-679-279-1	Sequence 1, Appli
27	159.5	4.0	3180	3 US-09-248-796A-20	Sequence 20, Appl
28	159.5	4.0	4647	3 US-09-252-991A-5730	Sequence 5730, Ap
29	159.5	4.0	10419	3 US-09-408-020-3	Sequence 3, Appli
30	159.5	4.0	42432	3 US-09-408-020-2	Sequence 2, Appli
c	159	3.9	13805	3 US-09-902-540-1083	Sequence 1083, Ap
32	157.5	3.9	2067	3 US-09-489-039A-2642	Sequence 2642, Ap
33	157	3.9	3155	3 US-09-710-279-3881	Sequence 3881, Ap
c	156.5	3.9	1860	3 US-09-252-991A-9781	Sequence 9781, Ap
c	156.5	3.9	5132	3 US-09-902-540-597	Sequence 597, App
36	156	3.9	26930	3 US-09-902-540-1228	Sequence 1228, Ap
37	156	3.9	72704	3 US-09-902-540-1273	Sequence 1273, Ap
c	155.5	3.9	2733	3 US-09-710-279-3107	Sequence 3107, Ap
39	155	3.8	2235	2 US-08-418-782-1	Sequence 1, Appli
40	155	3.8	2235	2 US-08-228-662-1	Sequence 1, Appli
41	155	3.8	2235	2 US-08-852-219-1	Sequence 1, Appli
42	154	3.8	3900	3 US-09-023-655-1420	Sequence 1420, Ap
c	154	3.8	16013	3 US-09-949-016-12988	Sequence 12988, A
44	153	3.8	2625	9 5457037-4	Patent No. 5457037
45	153	3.8	3050	3 US-09-710-279-3661	Sequence 3661, Ap

ALIGNMENTS

RESULT 1
US-09-533-559-7511
; Sequence 7511, Application US/09533559
; Patent No. 6902887
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjarke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE OF INVENTION: Expression
; FILE REFERENCE: 5849.200-US
; CURRENT APPLICATION NUMBER: US/09/533,559
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 09/273,623
; EARLIER FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7511
; LENGTH: 1103
; TYPE: DNA
; ORGANISM: Tricoderma reesei
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(1103)
; OTHER INFORMATION: n = A,T,C or G
US-09-533-559-7511
Alignment Scores: 1.47e-25 Length: 1103
Pred. No.: 392.50 Matches: 114
Score:

Percent Similarity:	45.0%	Conservative:	36
Best Local Similarity:	34.2%	Mismatches:	125
Query Match:	9.7%	Indels:	58
DB:	3	Gaps:	10

US-09-917-376-3 (1-740) x US-09-533-559-7511 (1-1103)

QY	302	GlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTrp	321
Db	3	GGCCVTGGCTCGATTGTCAAAGCCAGGAACCTTGTTGTTGCTTCTTGAACCTCTGG	62
QY	322	TrpProAspThrIleIlePheArgSerThrAspGlyGlyAlaThrTrpThrArgIleTrp	341
Db	63	TGGCCAGATGCTCAGCTGTTTCGTCACCGACTCTGGGACACATGAGAGCCGATCTGG	122
QY	342	AspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGluPro	361
Db	123	CGGTGGGCGGACTATCCGACTGACACCTATTACTACAGCATCTCACTCCCAAGCACCG	182
QY	362	TrpLeuThrPheGly---ValGlnProAsnProProValProSer-----Pro	376
Db	183	TGGATCAAGAACAACTTTATTCGATGTGACGAGCGAGTCACCGTCCGATGGTCINATCAAG	242
QY	377	LysLeuGlyTyrMetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeu	396
Db	243	CGCTCGGCTGGATGATTGAGTCTNTCGAGATTGACCCACCCGACAGCAANNACTGGCTT	302
QY	397	TyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuTyrLysTyrAspSer----	414
Db	303	TTACGGCACCGGAATGACATNTTGGCGGCCAGATTTCACCACTGGACACGGCGCC	362
QY	415	-----GlyGlyGlnIleHisIleAlaProMetValLys-GlyLeuGluGluThrAlav	432
Db	363	ACAATGTGGTCAATCCAATTACTTGGCAGCGGATTCGAAGAAAT-----TTTCCGT	416
QY	432	alaAsnAspLeuIleSerProSerGly-----	441
Db	417	TCAAGGACCTGGCCCTTTTACCGGGGGGAAGCGAGCTTTTGCCGCCAAGTCGGGAGCG	476
QY	442	-----AlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspV	459
Db	477	ANCAACGGGTTTACGTTTGGCCGAGAAACACGACCTTGG-----	517
QY	459	alThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThr-SerVal	478
Db	518	-----ACATTCGCGCAGAGCGTTTGGGCAACTCCACATGGGCCACTCGACGAGCGTC	572
QY	479	AspTyrAla--GluLeuAsnProSerIleIleValArgAlaGlySer-----	493
Db	573	GACTACGCGGGAATCGGTCAAGAGCGTTCGTCCGCGTCGGCAACACCGCGCGGACG	632
QY	494	-----PheAspProSerSerGlnProAsnAspArgHisValAlaPheS	508
Db	633	CAACMAGTGCCCATTTTTTCCGACGGCGGGCGCGAGCTNGGAA-----GCAATTC	686
QY	508	erThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGlyValThrThrGlyG	528
Db	687	GAACTAACGCT-----GNTCCGAAACACCGTTTCCATTGAAACGGCG	728
QY	528	lyThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaProGlyAspProGlyG	548
Db	729	CGCGGCTGGCTATTTCGCGCCGACGCGACACGATCTCTGTGTCGACCGCTCGTCCGGC-	787
QY	548	lnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProA	568
Db	788	-----GTGACGGCTCGCAGTTCCAGGCGAGCTTTCGCTCGGTCGACGCTCCCGC	839
QY	568	laAsnAlaGlnIleArgSerAspArgValAsnProLysThrPheTyrAlaLeuSerAsnG	588
Db	840	CGGGCGCGTCAATCGNCTCGGACAGAGAACCAACAGCGTNTTCTACGCGCGCTCGGAT	899
QY	588	lyThrPheTyrAspSerThrAspGlyGly	597

Df 900 CGACCTTTTACGTAGCAAGCACCGGC 928

RESULT 2
US-09-252-991A-13873/c
; Sequence 13873 Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13873
; LENGTH: 3129
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13873

Alignment Scores:
Pred. No.: 4,22e-10 Length: 3129
Score: 227.50 Matches: 218
Percent Similarity: 31.3% Conservative: 90
Best Local Similarity: 22.2% Mismatches: 360
Query Match: 5.6% Indels: 319
DB: Gaps: 52

US-09-917-376-3 (1-740) x US-09-252-991A-13873 (1-3129)

Qy 20 AspGlyIleValPheAsnGluGlyAla----ProGlyIleLeuTyrrValArgThrAspIle 38
Db 2993 AACGGCGTGTCATCAGCGGCACCAGCCGCCGCGCGTGCACCGTGCACCGTCCACCGATGCC 2934
Qy 39 GlyGlyMetTyrrArgTrpAspAlaAlaAsnGlyArgTriPileProLeuLeuAspTrpVal 58
Db 2933 GCC-----GGCAACCCGATAGGCGAGGTACCGCGCGAC 2901
Qy 59 GlyTrpAsnAsnTrpGlyTyr-----AsnGly-----ValVal 69
Db 2900 GGCAGCGGCAACTGGAGCTTCACCGCGGCACGCCCGCCGCAACGCGACGGTGATCGTC 2841
Qy 70 SerIleAlaAlaAspProIle-----AsnThrAsnLysValTrpAla-----Ala 84
Db 2840 GCACGCGCACCCAGCACCGCATACCGGCCCGCGCGCCGCGCCACCGTGGACGCG 2781
Qy 85 ValGlyMetTyrrThrAsnSerTrpAspProAsnAspGlyAlaIleLeu-----100
Db 2780 GTGCGCGCGCGCGCGGTGATCGATCCGAGCAACGCGACGACCATCAGCGGCACCGCG 2721
Qy 101 -----ArgSer 102
Db 2720 GAGCGCGCGCCAAGGTGATCTCACCGAGCGCAACCGCATCGCGGAACACCC 2661
Qy 103 SerAspGlnGlyAlaThrTrpGlnIleThrPro---LeuProPheLysLeuGly-----119
Db 2660 GCCGACGGCAGCGGCAACTGGACCTTCAGCGCCCGCCAGCCGCGTGGCCAAACGCGGTG 2601
Qy 120 -----GlyAsnMetProGlyArgGlyMetGlyGluArg 130
Db 2600 GTCAACCGCGTGCGCCGAGGACCTCGCGGCAATACCGCGCCCGCAGCGGCACACTACCGGTG 2541
Qy 131 LeuAlaValAspProAsn-----AsnAspAsnIleLeuTyrrPhe 143
Db 2540 GAGCGCGTGCGCGCGAACAACGCGCTGTGTTCAATCCGAGCAACGCGCAACTGTCTACCGGT 2481
Qy 144 GlyAlaProSetGlyLysGlyLeu-----Trp 152
Db 2480 ACCGCGCGCGCGCGAGCACCGTGACCTTGACCGAGCGGCAACGGCAACCGCATCGGCCAG 2421

Qy	153	ArgSerThrAspSerGlyAlaThrTrp	-----SerGlnMetThrAsn	167
Db	2420	ACCACGCCGATGGCAGCGCAACTGGAGCTTCACGCCCGCTCGCAACTACCAAC	--- 2364	
Qy	168	ProAspValGlyThrTyrlle	---AlaAsnProThrAspThrThrGlyTyrGlnSerAsp	186
Db	2363	-----GGCACCGTGGTCAACAGTGACCGCAGCAGCAGCCCGCGCAATACACGC	--- 2316	
Qy	187	IleGlnGlyValValTrpValAlaPheAspLysSerSerSerSerLeuGlyGln	----- 204	
Db	2315	-----GCTCCCGCTACCCAGCGGTGGATCTCTCGCTGCCGTGCATCCCGCAGGTGGAT	2262	
Qy	205	---AlaSerLysThrIlePheValGlyValAlaAspProAsnAsnProValPhe	----- 221	
Db	2261	CCGAGCAACGGTTCCGTGATCAGCGGCACCGCGGACCGCGCAACCATCATCATCACC	2202	
Qy	222	-----TrpSerArgAspGlyAlaThrTrpGlnAla	232	
Db	2201	GATGGCAACGGCAACCCGATTGGCCAGGTACACCGCCGACGGCAGCGGTAACTGGTCTTC	2142	
Qy	233	ValProGlyAlaPro	----- 237	
Db	2141	ACTCCAGGCATCCGCTGCCGATGGCAGCGTGCACGTGGTGGCGGCACGCCCAAGC	2082	
Qy	238	-----ThrGlyPheIleProHisLysGlyVal	246	
Db	2081	AATGTGCACAGTCGCGCGCGGTGATCACTGTGGATGGGTGGCGCGCGCGCGGTG	2022	
Qy	247	PheAspProValAsnHisValLeuTyrlleAlaThrSerAsnThrGlyGlyProTyr	--- 265	
Db	2021	ATCGATCCGAGCAGCGCACCCGAGATAGCGGTACCGCGGAGCGCGCGCGTATC	1962	
Qy	266	-----AspGlySerSerGlyAspValTrpLysPheSerValThr	---SerGlyThrTrp	282
Db	1961	CTACGATGGCGCGGCAACCCGATCGGCCAGCCACCGCGGAGCGCGCAACTGG	1902	
Qy	283	ThrArgIleSerProValPro	-----SerThrAsp	292
Db	1901	AGTTTACCCCGCGGCAACCCCGTGGCCAAACCGCACCGTGATCAACCGCTGGCGCCAGGAC	1842	
Qy	293	ThrAlaAsnAspTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisPro	----- 310	
Db	1841	CCGCGCGCAATACACAGCGTCCGCGCAGCGTCACCGTTCGATGCCATCGCCCGCGCG	1782	
Qy	311	-----AsnThrIleMetValAlaThrGlnIleSerTrpTrpProAsp	324	
Db	1781	CCGGTGATCAATCCGAGCAACGAGTGGTGCATCAGCGGTACGGCGGAGCCGGGGCCACG	1722	
Qy	325	ThrIleIlePheArgSerThrAspGlyGly	-----AlaThrTrpThr	338
Db	1721	GTGATCTC-----ACCGAGCGCAACGGCAACCCGATCGGCCAGGTCAACCGCGAC	1671	
Qy	339	ArgIleTrpAspTrpThr-SerTyrProAsnArg	-----SerLeu	351
Db	1670	GGCAGCGGAAGTGGGTTCACCGCCGCGCACCGCTGGCCAAATGGCAGGTGATCAAT	1611	
Qy	351	uArgTyValLeuAspIleSerAlaGluProTrpLeuThrPheGlyValGlnProAsnPr	371	
Db	1610	GCGTGG-----CCGAGGAGCGCGCGGCAACACAGCATC-----CC	1572	
Qy	371	oProValProSerProLysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAs	391	
Db	1571	ACCAGCGCCA---CCGTGACTCGCTGGCGCCAGCA-GCCCGGTGATCGATCCGAGCAA	1516	
Qy	391	nSerAspArgMetLeuTyrcGlyThr	-----GlyAlaThrLeuTyrcAlaThrAsnAs	408
Db	1515	CGGTAGC---GTGATCCCGGTACCCCGAGGCTGGTGCACCGGTGATCTCACC-----	1464	
Qy	408	pLeuThrLysTrpAspSerGlyGlnIleHisIleAlaProMetValLysGlyLeuGlu	428	
Db	1463	-----GACGGCAACGGCAAC-----CCGATCGGCGAGTCAACCGC	1429	

Qy	428	uGluThrAlaValAsnAspLeuIleSerProProSerGlyAlaProLeu	-----	444
Db	1428	CGATGGCAGCGCAACTGGAGCTTCACGCC	-----GGCACGCCGCTGCAATGGCAC	1375
Qy	445	-----IleSerAlaLeuGly	---AspLeuGlyGlyPheThrHisAlaAspVal---ThrAl	461
Db	1374	GGTGGTCAATCGGTGGCCAGAGACGCTGCGGCAACAC	CAGCGGCCCGCGCCAGCACCA	1315
Qy	461	aValProSerThrIlePheThrSerProValPhe	-----ThrThrGlyThrSerVal---	478
Db	1314	GGTGGACTCGGTGGCGCGCGCGCGGTGATCGACCG	AGCAACGGCAGCGTGATCGC	1255
Qy	479	AspTyrAlaGluLeuAsnProSerIleIleValArgAla	GlySerPheAspProSerse	498
Db	1254	CGGTACCGCGAAGCGGTGCAGCGGTGATCTCA	CCGATGGCGCGCAACCGCATCGG	1195
Qy	498	xGlnProAsnAspArgHisValAlaPheSerThrAsp	GlyGlyAsnTrp	515
Db	1194	CCAGGCC	-----ACCGCGATGGCAGCGCAACTGGAGCTTCAC	1156
Qy	516	-----	PheGlnGlySerGluProGlyG1	523
Db	1155	CCCGGGCACCGCGTGGCCAAACGCACGGTGATCA	TGTCGGTGGCCAGGATCGGCCCGG	1096
Qy	523	yValThrThrGlyGlyThr	-----	529
Db	1095	CAATACACAGCGCCGACACGACGACACCGTGG	ACCGTGGCCACCCCGGTGGT	1036
Qy	530	-----	ValAlaAlaSerAlaAspGlySerArgPheValTrpAl	542
Db	1035	CAACCCGAGCAACGGCAGCGTGATCGCGGTAC	CGGGAAACCGCGGCCACGGTGATCTCT	976
Qy	542	aProGlyAspProGlyGlnProVal	-----ValTyrAlaValGlyPheGlyAsnSerTr	560
Db	975	CACCGAGCGCGCGCAACCGATCGCGCCAGGTCA	CCGCCAGCGGACGCGCAAC	919
Qy	560	pAlaAlaSerGlnGlyValPro	---AlaAsnAlaGlnIleArgSer	574
Db	918	GAGCTTCACGCGCGGCACGCGCTGGCCAAACGG	TCGGTGATCAATCGCTGGCCACGA	859
Qy	575	-----	AspArgValAsnProIlysth	581
Db	858	CGCGCGCGCAACACACGCGCGCGCCAGCAC	CCGTCGGTACGATCGGTAGCCCGGCCAC	799
Qy	581	r---PheTyrAlaLeuSerAsnGlyThrPheTyrArg	Ser-----	593
Db	798	CCCGGTGCTCGATCCGAGCAACCGTACGCTGAT	CAGGTACGGTGACGTCGGTACCGCGGAGCC	739
Qy	594	-----ThrAspGlyGlyValThrPheGlnProVal	AlaAlaGlyLeu-ProSerS	610
Db	738	GGTGATCCTCAGCAGCGGCGC	-----GGCAACCCGATACGGCAGGCCACCCCGATGG	685
Qy	610	erGlyAlaValGlyValMetPheHisAlaValPro	GlyLysGluGlyAspLeuTrpLeuA	630
Db	684	CAGCGGCAACTGGAGCTTCACTCCGGGCACA	-CCGCTGACCAACGGCAGCGTATCAATG	626
Qy	630	laAlaSerSerGlyLeuTyrHisSerThrAsnGly	-----GlySerSerTrpSerAlaI	648
Db	625	CGGTGGCCAGGACCGCGCGCAACACGACGCGT	TCGGTCACACCCACAGTCGACGCGCGG	566
Qy	648	leThrGlyValSerSerAlaValAsnValGlyPhe	GlyLysSerAlaProGlySerSert	668
Db	565	TGGCCCGGCCACCCCGGTGATCGACCGAGCAAT	TGTTGTCAAACCTCAGCGGCACCGCGG	506
Qy	668	yrProAlaValPheValGlyThrIleGlyGly	-----	679
Db	505	AACCCGCGCTCCGGGTGATCTCACCAGATG	SGCAACCCCGATCGGCGACAGCCCTCG	446
Qy	680	-----	Valt	681
Db	445	CCGACGATGGGCAACTGGAACCTTCACACCG	GGCACCGCTGGCCAAACGGCACCGGTGG	386
Qy	681	hrGlyAlaTyrArgSerAspAspCysGlyThrThr	TrpTrpValLeuIleAsnAspAspGlnH	701

Db 385 TCACCGCGTGGCCAGGACCGCGCGCAATACCAG-----CGTCCGGCCAGC 336
 Qy 701 isGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArg-Val 720
 Db 335 ACCACGGTGGATACGGTGGCGCGCCACCGCGTGTATCAATGCCAGCAACGCGAGGTG 276
 Qy 721 TyrIleGlyThrAsn-----GlyArgGlyIleValTyrGlyAspIleGlyAlaPro 738
 Db 275 ATACCGCGGACCGCGAGGTGGCGGCAAAAGTGTCTCACCGACGCAACGCGCAACCCG 216
 Qy 739 SerGly 740
 Db 215 ATCGCG 210

RESULT 3.

US-09-252-991A-13656
 ; Sequence 13656, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 13656
 ; LENGTH: 8211
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-13656

Alignment Scores:

Pred. No.: 1.84e-09 Length: 8211
 Score: 227.50 Matches: 218
 Percent Similarity: 31.3% Conservative: 90
 Best Local Similarity: 22.2% Mismatches: 360
 Query Match: 5.6% Indels: 319
 DB: 3 Gaps: 52

US-09-917-376-3 (1-740) x US-09-252-991A-13656 (1-8211)

Qy 20 AspGlyIleValPheAsnGluGlyAla---ProGlyIleLeuTyrValArgThrAspIle 38
 Db 3574 AACGGCGTGTCTATCAGCGGCGACCGCGCGGCGGTGCCACCGTGCACCTTCACCGATGCC 3633
 Qy 39 GlyGlyMetTyrArgTrpAspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpVal 58
 Db 3634 GGC-----GGCAACCGGATAGGCGAGGTCCACCGCGCAG 3666
 Qy 59 GlyTrpAsnAsnTrpGlyTyr-----AsnGly-----ValVal 69
 Db 3667 GGCAGCGGCACTGGAGCTTCACCGCGGCGACCGCGCGCGCGCAACGCGCGGTGTATGTC 3726
 Qy 70 SerIleAlaAlaAspProIle-----AsnThrAsnLysValTrpAla-----Ala 84
 Db 3727 GCCACGGCCACCGACCGCGCGCAATACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3786
 Qy 85 ValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeu-----100
 Db 3787 GTGGCGCGCGCGCGCGGTGATCGATCCGAGCAACGCGACGACCATCAGCGGCAACCGCG 3846
 Qy 101 -----ArgSer 102
 Db 3847 GAGCGCGGCGCGNAGTGTATCTCACCGCGGCAACCGCGCAACCGCGATCGCGGCAACCGC 3906
 Qy 103 SerAspGlnGlyAlaThrTrpGlnIleThrPro---LeuProPheLysLeuGly-----119
 Db 4957 GCGGTGG-----CCGAGCGCGCGCGGCAACACAGCAGTCT-----CC 4995

Db 3907 GCCACGGCAGCGCAACTGCACCTTACGGCGCGCCACCGCCCTGGCCACCGCAGCGTG 3966
 Qy 120 -----GlyAsnMetProGlyArgGlyMetGlyGluArg 130
 Db 3967 GTCAACGCGGTGGCCAGGACCTTGGCGGCAATACCGGCGCGCGCAGCAGCACTACCGTG 4026
 Qy 131 LeuAlaValAspProAsn-----AsnAspAsnIleLeuTyrPhe 143
 Db 4027 GACCGGTGGCGCGCAACACCGCTGTGTCAATCCGAGCAACCGCAACCTGCTCAACGCT 4086
 Qy 144 GlyAlaProSerGlyLysGlyLeu-----Trp 152
 Db 4087 ACCCGCGAGCGCGCAGCACCCTGACCTTGACCGGACGCGCAACCGCGATCGGCCAG 4146
 Qy 153 ArgSerThrAspSerGlyAlaThrTrp-----SerGlnMetThrAsnPhe 167
 Db 4147 ACCACCGCGATGCGAGCGGCAACTGGAGCTTCACGCGCGGTCTCGCACTACCCACAC 4203
 Qy 168 ProAspValGlyThrTyrIle---AlaAsnProThrAspThrGlyTyrGlnSerAsp 186
 Db 4204 -----GGCACCGTGTCAACGTGACCGCGGAGCGCGCGCGCAATACCCAGC 4251
 Qy 187 IleGlnGlyValValTrpValAlaPheAspLysSerSerSerSerLeuGlyGln-----204
 Db 4252 -----GCTCCCGCTACACGACCGGTGATCTCTGCTGCCGTGATCCCGCAGGTGGAT 4305
 Qy 205 ---AlaSerLysThrIlePheValGlyValAlaAspProAsnAsnProValPhe-----221
 Db 4306 CCGAGCAACGGTTCGGTGTATCAGCGGACCGCGGAGCGCGGCAACCATCATCATCACC 4365
 Qy 222 -----TrpSerArgAspGlyGlyAlaThrTrpGlnAla 232
 Db 4366 GATGGCAACGCGCAACCGATGGCCAGGTACCGCGCGACGCGGCGGTAACTGGTCTTC 4425
 Qy 233 ValProGlyAlaPro-----237
 Db 4426 ACTCCAGGCATCCCGCTGCCCGGATGGCACGGTGTGTCAACGTGGCGCGCAGCCCAAGC 4485
 Qy 238 -----ThrGlyPheIleProHisLysGlyVal 246
 Db 4486 AATGTCACAGTGGCGCGCGGTGATCACTGTGGATGGTGGCCCGCGCGCGCGCGGTG 4545
 Qy 247 PheAspProValAsnHisValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyr---265
 Db 4546 ATCGATCCGAGCAACGCGCACCGAGATAGCGGTACCGCGGAGCGCGCGCGCGCGGTATC 4605
 Qy 266 -----AspGlySerSerGlyAspValTrpLysPheSerValThr---SerGlyThrTrp 282
 Db 4606 CTCACCGATGGCGCGCGCAACCGATCGCGCGCGCACCGCGCGCGCGCGCGCACTGG 4665
 Qy 283 ThrArgIleSerProValPro-----SerThrAsp 292
 Db 4666 ACGTTACCCCGCGCGCACCGCGTGGCCAAACCGCACCGGTGATCAACCGCGTGGCCAGGAC 4725
 Qy 293 ThrAlaAsnAspTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisPro-----310
 Db 4726 CCGCGCGCAATACCGAGGTTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4785
 Qy 311 -----AsnThrIleMetValAlaThrGlnIleSerTrpTrpProAsp 324
 Db 4786 CCGGTGATCAATCCGAGCAACGGAGTCTGTATCAGCGGTACGCGGAAAGCGCGGCGCGCAG 4845
 Qy 325 ThrIleIlePheArgSerThrAspGlyGly-----AlaThrTrpThr 338
 Db 4846 GTGATCTCTC-----ACCGACGGCAACCGCAACCGCATCGCGCGCGCGCGCGCGCG 4896
 Qy 339 ArgIleTrpAspTrpThr-SerTyrProAsnArg-----SerLe 351
 Db 4897 GGCAGCGCGAAGTGGGCTTTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4956
 Qy 351 uArgTyrValLeuAspIleSerAlaGluProTrpLeuThrPheGlyValGlnProAsnPr 371
 Db 4957 GCGGTGG-----CCGAGCGCGCGCGGCAACACAGCAGTCT-----CC 4995

Qy	371	oProValProSerProLysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAs	391
Db	4996	ACCAAGCCCA---CCGTCGACTCGTGGCCAGCA-GCCCGGTGATCGATCCGAGCAA	5051
Qy	391	nSerAspArgMetLeuTyRglyThr-----GlyAlaThrLeuTyAlaThrAsnAs	408
Db	5052	CGGTAGC---GTGATCGCGGTACCGCCGAGGCTGGTCCACGCTATCCTCACC----	5103
Qy	408	pLeuThrLysTrpAspSerGlyGlyGlnIleHisIleAlaProMetValIscyLeuGl	428
Db	5104	-----GAGGGCAACGGCAAC-----CGATCGGCAGGTCAACCGC	5138
Qy	428	uGluThrAlaValAsnAspLeuIleSerProProSerGlyAlaProLeu-----	444
Db	5139	CGATGGCAGCGGCAACTGGAGCTTACGCC-----GGCAGCCGCTGTCATCGCAC	5192
Qy	445	-----IleSerAlaLeuGly---AspLeuGlyGlyPheThrHisAlaAspVal---ThrAl	461
Db	5193	GGTGGTCAATGCGGTGGCCAGGACGCTGCCGCAACACACGAGCGGCCGCCAGCAC	5252
Qy	461	aValProSerThrIlePheThrSerProValPhe-----ThrThrGlyThrSerVal--	478
Db	5253	GGTGGACTCGTGGCGCGCGCGCGCGGTGATCGACCCGAGCAACGCGAGCGTGATCGC	5312
Qy	479	AspTyAlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspProSerSe	498
Db	5313	CGGTACCGCGGAAGCCGTGCGAGGTGATCTCTCACCGATGGCGCGCAACCCGATCGG	5372
Qy	498	rGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLyAsnTrp-----	515
Db	5373	CCAGGCC-----ACCGCGATGGCAGCGCACTGGAGTTCTCAC	5411
Qy	516	-----PheGlnGlySerGlyProGlyGl	523
Db	5412	CCCGGCGACGCGCTGGCCAAACGCGAGGTGATCAATGCGGTGGCCAGGATCCGCCGG	5471
Qy	523	yValThrThrGlyGlyThr-----	529
Db	5472	CAATACGAGCGGCCGACCAACGACACACCGGTGGACGGTGGTGGCCGCCACCCGGTGT	5531
Qy	530	-----ValAlaAlaSerAlaAspGlySerArgPheValTrpAl	542
Db	5532	CAACCCGAGCAACGCGAGCGTGTATCCCGGTACCCGCGAAGCGCGCCACGCTGTACTCT	5591
Qy	542	aProGlyAspProGlyGlnProVal-----ValTyAlaValGlyPheGlyAsnSerTr	560
Db	5592	CACGACGGCGGGCGCAACCCGATCGCCNGTCAACGCGCGACGCGCGGCAC---TG	5648
Qy	560	pAlaAlaSerGlnGlyValPro---AlaAsnAlaGlnIleArgSer-----	574
Db	5649	GAGTTTACGCCCGGACGCGCGTGGCCCAACGCGTCTGGTATCAATGCTGCGTCCGACGA	5708
Qy	575	-----AspArgValAsnProLysTh	581
Db	5709	CGCGCGCGGCAACACCAACGCGCGCCGCCAGCACACCGGTGACTCGTAGCCCGGCCAC	5768
Qy	581	r---PheTyAlaLeuSerAsnGlyThrPheTyArgSer-----	593
Db	5769	CCCGGTGCTCGATCCGAGCAACGGTATACGTGTATCAGCGGTACCGCGAAGCCGGGCCAC	5828
Qy	594	-----ThrAspGlyGlyValThrPheGlnProValAlaAlaGlyLeu-ProSerS	610
Db	5829	GGTGTATCTTACCAGCGGGCC-----GGCAACCCGATACGGCAGGCCACCGCCGATGG	5882
Qy	610	erGlyAlaValGlyValMetPheHisAlaValProGlyLysGluGlyAspLeuTrpLeuA	630
Db	5883	CAGCGCAACTGGAGTTTCACTCCGGGCACA-CCGCTGACCAACGCGCATCAATG	5941
Qy	630	lAlaSerSerGlyLeuTyHisSerThrAsnGly-----GlySerSerTrpSerAlar	648
Db	5942	CGGTGGCCAGGACGCCCGCGGCAACACCAAGCGGTCCGGTTCAGCACCACTGAGCGCG	6001

Qy	648	leThrGlyValSerSerAlaValAsnValGlyPheGlyLysSerAlaProGlySerSert	668
Db	6002	TGGCCCGGCGACCCCGGTGATCGACCGGAGCAATGTTGTCAAACTCAGCGGCACGCGC	6061
Qy	668	yrProAlaValPheValValGlyThrIleGlyGly-	679
Db	6062	AACCCGCGTCCGGGTGATCCTCACCAGATGGCAATGGCAACCCGATCGGCACACCTCG	6121
Qy	680	-----Valt	681
Db	6122	CCGACGTAGCGGCACTGGACCTTCACACCGGGCAGCGCGTGGCCACCGCACGGTGG	6181
Qy	681	hrGlyAlaTyraGserAspAspCysGlyThrThrTrpValLeuIleAsnAspGlnH	701
Db	6182	TCAACGCGGTGGCCAGACCCGCGCGCAATACCAAG-----CGTCCGGCCAGC	6231
Qy	701	isGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgArg-Val	720
Db	6232	ACCACGGTGGATACGGTGGCGCGGCACCCCGGTGATCAATCCAGCAACGGCAGCGGTG	6299
Qy	721	TyrIleGlyThrAsn-----GlyArgGlyIleValTyrglyAspIleGlyGlyAlaPro	738
Db	6292	ATACCGGCACCGCGGAGGTGGGGCCAAAGTGATCTCACCGCGGCAACGCAACCCG	6351
Qy	739	SerGly 740	
Db	6352	ATCGGC 6357	
RESULT 4			
US-09-252-991A-13774/c			
; Sequence 13774, Application US/09252991A			
; Patent No. 6551795			
; GENERAL INFORMATION:			
; APPLICANT: Marc J. Rubenfield et al.			
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUD			
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS			
; FILE REFERENCE: 107196.136			
; CURRENT APPLICATION NUMBER: US/09/252, 991A			
; CURRENT FILING DATE: 1999-02-18			
; PRIOR APPLICATION NUMBER: US 60/074, 788			
; PRIOR FILING DATE: 1998-02-18			
; PRIOR APPLICATION NUMBER: US 60/094, 190			
; PRIOR FILING DATE: 1998-07-27			
; NUMBER OF SEQ ID NOS: 33142			
; SEQ ID NO 13774			
; LENGTH: 4188			
; TYPE: DNA			
; ORGANISM: Pseudomonas aeruginosa			
US-09-252-991A-13774			
Alignment Scores:			
Pred. No.: 1.02e-07 Length: 4188			
Score: 203.00 Matches: 186			
Percent Similarity: 32.4% Conservative: 85			
Best Local Similarity: 22.2% Mismatches: 297			
Query Match: 5.0% Indels: 274			
DB: 3 Gaps: 46			
US-09-917-376-3 (1-740) x US-09-252-991A-13774 (1-4188)			
Qy	47	AlaAsnGlyArgTrp-----IleProLeuLeuAspTrpValGlyTrpAsn	61
Db	2496	GCCAAACGGCAACTGGTCTTTTCACGCGTCCACCCCGCTGCCGAGC-GTACCGTGTCA	2439
Qy	62	AsnTrp-----GlyTyraGlyValValSerIleAlaAlaAspProIleAsnThrAsn	79
Db	2438	ACGTGTGGCGCAGGAGTACCGCGGCAACAGCAGTCCGCGGCGCAGCGTTA-----	2388
Qy	80	LysValTrpAlaAlaValGlyMetTyThrAsnSerTrpAspProAsnAspGlyAlaIle	99
Db	2387	---CCGTGGAT-GCCCGTGGCGCGCGGCACGCCACCGTCGATCCGAGCAACGATACGACC	2332
Qy	100	LeuArg---SerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeu	118

Db 2331 CTACGGCGACCGCGCGCGCGCTACCGTGACCGCTGACC----- 2290
Qy 119 GlyGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAsp 138
Db 2289 GACGGCAAC-----GGCAACCCGATTGGCCAGGTACCGGCC----- 2254
Qy 139 AsnIleLeuTyrPheGlyAlaProSerGlyGlyLeuTyrArgSerThrAspSerGly 158
Db 2253 -----GACGGCAGCGGCAACTGG----- 2236
Qy 159 AlaThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyrIle---AlaAsnPro 177
Db 2235 ---ACCTTACCGCGAGCAGCGCGCTGCCCAAC---GGCAGCGGTGTCACCGCACCGCGCT 2182
Qy 178 ThrAspThrThrGlyTyrGlnSerAspIleGlnGlyValValTyrValAlaPheAspLys 197
Db 2181 ACCGACCGGTCCGCAACCGCGAGTTCGCGCGCAGCGTCAACCGTGGAGCCCGTGGCCACCG 2122
Qy 198 SerSerSerSerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspPro--- 216
Db 2121 GCCACGGCAGTGGTCAACCGCGAGCAACCGCACCGCTCAGCGCGCACCGCGCGCGCGC 2062
Qy 217 -----AsnAsnProValPheTrp-----SerArgAspGly 226
Db 2061 GCCACCGTGACCGTGGCGATGGCAACCGGCAATCCCATCGCGGCAAGGTCAACCGCGATGGC 2002
Qy 227 GlyAlaThrTrpGlnAlaValProGlyAlaPro----- 237
Db 2001 AGCGGCAACTGGAGCTTCACTCCGACCAACCGCGGTGGCCCAACGGCACCGGTGTCACGCC 1942
Qy 238 -----ThrGlyPhe 240
Db 1941 ACGGCCACCGAGCGCTCCGGCAACACACAGTGGCGGCAGCAGTGTACCGTGGATCGGTA 1882
Qy 241 IleProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsn 260
Db 1881 GCCCGCGCCACCGCAGTGATCAACCCGAGCAACCGCACCGCTCAGCGCGCACCGCGCGAG 1822
Qy 261 ThrGlyGlyProTyr-----AspGlySerSerGlyAspValTyrLysPheSerVal 277
Db 1821 CCGGGCAGCAGCGTACCTGACCGATGGCAACGGCAACCGATCGCGCCAGGTCAACCGCC 1762
Qy 278 Thr-----SerGlyThrTrpThrArgIleSerProValPro----- 289
Db 1761 GACGGCAGCGCACTGGAGCTTCAACCGCTCCACCGCTGCGCGGTGGCGGATGGACCGTGGTC 1702
Qy 290 -----SerThrAspThrAlaAsnAspTyrPheGlyTyrSerGlyLeuThrIleAsp 306
Db 1701 AACGCCACCGCGCCAGTCCGCGGGCAACACACAGCGCGCCAGGCGCAGCACCCCGTGCAT 1642
Qy 307 ArgGlnHisProAsnThrIleMetVal-----AlaThrGlnIleSerTrp 322
Db 1641 GGGGTGGCGCGCACCGCGCGCGCTCAACCTGTAGCAACGGCAGCAGCTCAGCGGCACT 1582
Qy 323 ProAsp-----ThrIleIlePheArg----- 329
Db 1581 CGGGAACCGGGCAGCAGGTGATCTCACCGAGGCAACCGCAATCCGATCGCGCGAGTGC 1522
Qy 330 SerThrAspGlyGlyAlaThrTrpThrArgIleTrpAspThrTrpSerTyrProAsnArg 349
Db 1521 ACCGCCAGCGCAGCGCAACTGGACC-----TACACCCCGT---CCACGCGGA 1476
Qy 350 SerLeuArgTyrValLeuAspIleSerAlaGluProTrpLeuThrPheGlyValGlnPro 369
Db 1475 TCG-----CCAACGGCACCGTGGTCAACGTGGCGCCAGCG 1437
Qy 370 AsnProProValPro-----SerProLysLeuGlyTrpMetAspGluAlaMetAla--- 386
Db 1436 CCGCGCGCAATAGCAGCGCGCGCGAGCTACCGGTGAGAC---TCGCAAGCCCGCGCGCT 1378
Qy 387 -----IleAspProPheAsnSerAspArgMetLeuTyrGlyThr-----GlyVala 401
Db 392

Db 1377 CCGGTGTCTAACCCGAGCAACCGGC---ACCACGTCTACGGCGCACCGCGAGCGCGCGCT 1321
Qy 402 ThrLeuTyrAlaThrAsn-----AspLeuThrLysTrpAspSer 414
Db 1320 ACCGTGAGCTGACGCGACCGCAACCGGATTTGGCCAGGTACCGCC---GACGGC 1264
Qy 415 GlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluGluThrAlaValAlaAsp 434
Db 1263 AGTGGCACTGAGCTTCAACACCGCGCACCGCTGGCCACCGCACCGTGTCAACGCC 1204
Qy 435 LeuIleSerProProSerGly-----AlaProLeuIleSerAlaLeuGlyAspLeu 451
Db 1203 ACGGCCAGCGACCGCGCAATACCGCGCTCCGGCCAGCACACCGCTGACGCTGGT 1144
Qy 452 GlyGlyPheThrHisAlaAspValThrAlaValProSerThrIlePheThrSerProVal 471
Db 1143 -----GCGCGCGCGCGCTGGTCAATCCGAGCAACGGAGTC 1105
Qy 472 PheThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleValArgAla 491
Db 1104 GTCATCAGCGCACCC-----GCCGAACCGGGCGCCACCGTGACCTGACCGAT 1057
Qy 492 GlySerPheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGly 511
Db 1056 GGCAGCGCAATCCGATCGGGCAG-----GTCACCGCGCGCACCGC 1018
Qy 512 GlyLysAsnTrp-----Phe 516
Db 1017 AGCGGCAACTGAGCTTCAACCGCTCCAGCGGTGGCGGATGGAAACCGTGTCAACGCC 958
Qy 517 GlnGlySerGluProGlyValThrThrGlyGly-----ThrValAlaAla 532
Db 957 ACCGTACCGACCGCGCGGC---AATACCGCGCGCCAGGCGCAGCACTACCGTGGACGCC 901
Qy 533 SerAla-----AspGlySerArgPheValTrpAlaPro 543
Db 900 ATCCGCGCGCGCACCGCGCACCGTCAACCTGAGCAATGGCAGCAGC---CTCAGCGCACT 844
Qy 544 GlyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsn----- 558
Db 843 CGGGAACCGGCGACCGGTGATCTCTACCGACGCAACCGCAATCCGATCGCGCGAGTGC 784
Qy 559 -----SerTrp----- 560
Db 783 ACCGCCAGCGCAGCGGCACTGAGCACTACACCGCTCCACCGCGATCGCCAACCGTACT 724
Qy 561 -----AlaIleSerGlnGlyValProAlaAsnAlaGln 571
Db 723 GTGCTCAACGTGTGGCGGAGGACCGCGCGGTAAACAGCAGCGCGCGCGCGCGTACC 664
Qy 572 IleArgSerAspArg-----ValAsnProLysThrPheTyrAlaLeuSer 586
Db 663 GTCGATTCCAGCGCGCGCGCGCGCGCGTGTATCAACCG-----AGC 622
Qy 587 AsnGlyThrPheTyrArgSerThrAspGlyValThrPheGlnProValAlaAlaGly 606
Db 621 AACGGCGTC-----GTCTACGCGGC 601
Qy 607 LeuProSerSerGlyAlaValGlyValMetPheHisAlaValProGlyLysGluGlyAsp 626
Db 600 ACCGCCGAGCGCGTGCACCGTGCACCTCACCAGTCCCGCGGCAACCCGATAGGCGAG 541
Qy 627 LeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsnGlyGlySerSerTrpSer 646
Db 540 GTC-----ACCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 511
Qy 647 AlaIleThrGlyValSerSerAla-----ValAsnValGlyPheGlyLysSerAla 663
Db 510 TTCACGGCGCGCACCG 451
Qy 664 ProGlySerSerTyrProAlaValPheValValGlyThrIleGlyValThrGlyAla 683
Db 450 ACCGGCAATACCG 392


```
Db 2619 GAGCCGCCGCGCAACACGAGCGCGCCGAGCAGCAGCGTGGACTCGGTAGCCCGCGCC 2560
Qy SerProValPhe-----ThrThrGlyThrSerValAsp---TyrAlaGluLeuAsnPro 485
Db 2559 ACCCGGTGTCTGATCCGAGCAGCGTACCGTATCAGCGGTACCGCGGAGCGCGGCC 2500
Qy SerIlelleValArgAlaGlySerPheAspProSerSerGlnProAsnAspArgHisVal 505
Db 2499 ACGGTATCTCAGCGAGCGCGCGGCAACCCGATACCGCAGGCC-----2455
Qy 506 AlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGlyValThr 525
Db 2454 -----ACCGCCATGTCAGCGCGCACTGG-----AGCTTCACTCCGGGACACCGCTG 2407
Qy 526 ThrGlyGlyThrVal-----AlaAlaSerAlaAspGlySerArgPheValTrpAlaPro 543
Db 2406 ACCAAGCGCAGGTGATCAATGCGGTGGCCGAGGACCGCCGCGCAACACCGGTCGC 2347
Qy 543 -----543
Db 2346 GTCAGCACCACAGTGGAGCGGTGGCCCGCCGACCCCGGTGATCGACCGAGCAATGGT 2287
Qy 544 -----GlyAspProGlyGlnProValValTyrAlaValGlyPheGly 557
Db 2286 GTCAAACTCAGCGCACCGCGCAACCGCGGTCCGGTGATCTCTCACCGATGGCAATGGC 2227
Qy 558 Asn-----SerTrpAlaAlaSerGlnGlyVal 566
Db 2226 AACCGATCGGCAGACCCCTCGCGAGCGGTAGCGGCAACTGGACCTTCACACCGGCGAGC 2167
Qy 567 ProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThrPheTyrAlaLeuSer 586
Db 2166 CCG-----CTGGCC 2158
Qy 587 AsnGlyThrPheTyrArg-----SerThrAspGly-----596
Db 2157 AACGGCAGCGTGTCAACCGCGTGGCCGAGGACCGCGCGCAATACCGCGTCCGCC 2098
Qy 597 GlyValThrPheGlnProValAlaAlaGlyLeuPro-----SerSerGlyAla 612
Db 2097 AGCACCAGCGTGATACGGTGGCGCGCGCCGCGGATCAATCGACCAACGCGCAGC 2038
Qy 613 ValGlyValMetPheHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSer 632
Db 2037 GTG---ATCACCAGCGCGCGAGTCCGCGCGCAAGTGATCTCTACCGAGCGCAACGGC 1981
Qy 633 SerGlyLeuTyrHisSerThr---AsnGlyGlySerSerTrpSerAlaIleThrGlyVal 651
Db 1980 AACCCGATCGCGAGACCGCGCGCGCGCGAGTGGCACTGGACCTTCACCCCGCGCAGC 1921
Qy 652 -----SerSerAlaValAsnValGlyPheGlyLysSerAlaProGlySerSer 667
Db 1920 CGCTCGCCCAACGGTACGGTGATCAACCGCGCGCGCGCGAGACCGCGCGGCGCAGCAGC 1861
Qy 668 TyrProAlaValPheValGlyThrIleGlyGlyValThrGlyAlaValArgSer 686
Db 1860 GGTCCGGCC-----AGCACAC-GGTGGACTCGGTGGCGCGCTCCCTCC 1817
```

RESULT 6

```
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: TUBERCULOSIS
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
```

```
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1

Alignment Scores:
Pred. No.: 0.0402 Length: 4411529
Score: 192.00 Matches: 194
Percent Similarity: 35.0% Conservative: 94
Best Local Similarity: 23.6% Mismatches: 323
Query Match: 4.8% Indels: 213
DB: 3 Gaps: 43

US-09-917-376-3 (1-740) x US-09-103-840A-1 (1-4411529)

Qy 10 AsnValAlaIleGlyGlyGlyPheValAspGlyIleValPheAsnGluGly-----27
Db 427279 AACCGAACATCGCGCGCGCAACATCGCGAC-----TTTAACTCGGATCGCA 427229
Qy 28 -----AlaProGlyIleLeuTyrValArgThrAspIleGly-----GlyMet 41
Db 427228 AACACCGGTCCGGGCTAACCGCGCTGTCAACAACATCGGTATCGCAACACCGCAAC 427169
Qy 42 TyrArgTrpAspAlaAlaAsn---GlyArgTrpIleProLeuLeuAspTrpValGlyTrp 60
Db 427168 TACAACATCGGTGTGCGCAACACCGGTAACTACAAAC-----ATCGGCTTC 427124
Qy 61 AsnAsnTrpGlyTyrAsnGlyVal---ValSerIleAlaAlaAsp-----74
Db 427123 GGCACACCGCGCAACACATCGCATCGCGCTGTCCGCGCAACACGATCGGTTTC 427064
Qy 75 ---ProIleAsnThrAsnLysValTrpAlaValGlyMetTyrThrAsnSerTrpAsp 93
Db 427063 GGCCCGCTGAACCGCGCGC-----ATCGCAACATGGCGCTTTC---AACCTGGCGAC 427013
Qy 94 ProAsnAspGlyAlaIleLeuArgSerSerAspGlnGlyAlaThrTrpGlnIleThrPro 113
Db 427012 AACAACTTTGGC-----427001
Qy 114 LeuProPheLysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaVal 133
Db 427000 -----ATGCCAACCGCGCAACTTCAACAGCGCATTTGCC 426965
Qy 134 AspProAsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLys---GlyLeuTrp 152
Db 426964 AACACCGCGCAACACATCGGTCTTGTCAACACCGCGCAACACACATCGCATCTGG 426905
Qy 153 ArgSerThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThr 172
Db 426904 CTGACCGCGCAGCGCTTGTCTCGGTTCAGCTCCCTGAACTCCGGCGCGCAACACCGGT 426845
Qy 173 TyrIleAlaAsnProThrAspThrGlyTyrGlnSerAspIleGlnGlyValValTrp 192
Db 426844 TTCTTCAACTCCGGCACCGCAACACCGGC-----426815
Qy 193 ValAlaPheAspLysSerSerSerSerLeuGlyGln---AlaSerLysThrIlePheVal 211
Db 426814 ---TTGTTCACTCCGCGCACCGCAACACCGCTTGTTCACCTCGGCGCGCAACGTC 426758
Qy 212 GlyValAlaAspProAsnAsnProValPheTrpSerArgAspGlyGlyAlaThrTrpGln 231
Db 426757 GGCATCGCAACATCGGCGCGGCTTCGCGCTCGCGCTTCCGCGCGCAACAGGAGTG 426698
Qy 232 AlaValProGlyAlaProThrGlyPheIleProHisLysGlyValPheAspProValAsn 251
Db 426697 GGCATCGCGCGCACCAACTCGGCGAGTTTC---AACATCGCGCTTGTTCACCTCGGCGCAC 426641
Qy 252 HisValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyrAspGlySerGlyAs 271
Db 271 -----
```

Db 426640 GGCAATGTGGCATCGGCAACTCGGGCACCGCAACGTCGGCATCGGCAACACACCGGCACC 426581
Qy 271 pValTrpLysPheSerValThrSerGlyThr-----Tr 282
Db 426580 GGCAACACCGGCATCGGAAACAGCGCAACTACAAACCGCGTCTCTCAACCGCGGCCTG 426521
Qy 282 p-----ThrArgLysProValProSerThrAspThrAlaAsnAs 296
Db 426520 GTCAACACCGGCATCGGCAACCGCGGCAACCAACACCGCGCTTTCAACATCGGCACC 426461
Qy 296 pTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAl 316
Db 426460 TTCACACCGGCATCGCAACCGCGGCACCTACAAACCGGCTCTACAAACCGGTAGC 426401
Qy 316 aThrGlnIleSerTrp-----TrpProAspThrIleIle-----PheArgSerThr----- 331
Db 426400 TACAACACCGGCATCGGCAACCGCGGAGACTACGCGACCGCGCGCTTCATCACCGGCAGC 426341
Qy 332 -----AspGlyGlyAlaThrTrpThrArgIleTrpAspTrpThrSer-- 345
Db 426340 ATGAACACCGCTTGTCTGTGGCGCGCCGACCGGAGCGCTCTGTGGCGGCCAACTACACC 426281
Qy 346 -----TyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGl 360
Db 426280 ATCACCATCGAGCGACCTGCGCGGTCTCTCAATGTGCATCCCGGTCAACATCCCATC 426221
Qy 360 uProTrpLeuThrPheGlyValGlnProAsnProValProSerProLysLeuGlyTr 380
Db 426220 ACCG-----CGGACATCACCAATGTCTCCATCCCGCCATTACGTTCCTCCAGCA-ATC----- 426170
Qy 380 pMetAspGluAlaMetAlaIleAsp-----ProPheAsnSerAspArgMetLeuTyrGlyTh 399
Db 426169 ----GAGCGCAGCGGAGCGTCGACATAGGCATCTCAGTGGCACCGCTTGGCCCGCT 426114
Qy 399 rGly-----AlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlnI 418
Db 426113 CGTCCGATCACTGTCATGCGGGGAGCGCTGCGCCCGCTGGACACCCATCGAAT 426054
Qy 418 eHisIleAlaProMetValLysGlyLeuGluThrAlaValAsnAspLeuIleSerPr 438
Db 426053 TGACTTCGGCC-----TCGCGCGGATCAACCTCAACATCGGCA 426012
Qy 438 oProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyPheThrHisAlaAs 458
Db 426011 GCCCGAGCGCTCCACCGTGATCAACATCGTGGCGGCGCGC----- 425969
Qy 458 pValThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSerVa 478
Db 425968 -----GCCGGCGCGATC-----AGCAT 425952
Qy 478 lAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspProSerSe 498
Db 425951 TCCGATCATCGACTTGGCGCAGCG-----CCCGGCTCTTCAACCGCCACAC 425904
Qy 498 rGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGl 518
Db 425903 CGGCCG-----TCGTGGGCTCTCTCAACTGGGTGTGG 425868
Qy 518 ySerGluProGlyGlyValThrThrGlyGlyThr-----ValAlaAl 532
Db 425867 CAGCGCATCGGCTGTGTGAATCTCGGCAACAACTCGGGCCTCTACAACTTCGCCACTAG 425808
Qy 532 aSerAlaAspGlySerArgPheVal-----TrpAlaProGl 544
Db 425807 CAGCATGGGAAATTCGGGCTTCCAAAATACTATGGGTGCTGCAGTGGGCTGGCG----- 425753
Qy 544 yAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSer-----Tr 560
Db 425752 -----AATTGGGCAACAGCATCTCGGCATCTA 425724
Qy 560 pAlaAserGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLy 580
Db 425723 CAACACCGGCTTGGGAGCACCGGCAATGTC-----TCGGGCTTGTCTCAACATCGG 425673

Qy 580 sThrPheTyrAla-----LeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyVa 598
Db 425672 CACCAACCTGGTGGTGGTTCAGAACCGGCGC----- 425639
Qy 598 lThrPheGlnProValAlaAlaGlyLeu-ProSerSerGlyAlaValGlyValMetPhe- 617
Db 425638 -ACCGAGACGACTTCAGCGTGGCTTGGCCAACTCGGTTCT---GGAACTCGGGTAG 425583
Qy 618 -----HisAlaValProGlyLysGluGlyAspLeuTrpLeuA 630
Db 425582 CGCAACATCGGCAACTACAACTGGGACGCCCAACATCGGCTCTACAACTGGGCAG 425523
Qy 630 la-----AlaSerSerGlyLeuTyrHisSerThrAsnGlyGlySerSerTrpSerA 647
Db 425522 CGCCAACTCGGCGACTTCAACTGGGACGCGCAACATCGGCGACTTCAACCTGGGCAG 425463
Qy 647 laIleThrGlyValSerSerAlaValAsnValGlyPheGlyLysSerAlaProGlySerS 667
Db 425462 CGCCAACTCGGCGACTTCAACTGGGACGCGCAACATCGGCTCTCGGCAACCGGAAACGGAACA 425362
Qy 667 eTyrProAlaValPheValValGlyThrIleGly---GlyValThrGlyValaTyrArgS 686
Db 425413 -----CTGACGGGCGCATCGGCAACATCGGCTTCGGCAACCGGAAACGGAACA 425302
Qy 686 eTyrProAlaValPheValValGlyThrIleGly---GlyValThrGlyValaTyrArgS 706
Db 425361 TCGGCATCGGCAATACCGGACCGCAACATCGGCTTCGGCAACCGGAAACGGAACA 425302
Qy 706 rPlyGlnAlaIleThrGlyAsp---HisAlaAsnLeuArgValTyrIle-GlyThrA 725
Db 425301 TCGGCATCGGCTACCGGCGACACATCGGCTTCGGCGCTCGGAACTCGGCGACCG 425242
Qy 725 snGlyArgGlyIleVal-----TyrGlyAspIleGlyGlyAlaProSerGly 740
Db 425241 GCAACATCGGCTATTCAACTCGGACCGGCAACATCGGCTTCGGCAACTCGGC 425186

RESULT 7

US-09-902-540-1027/c
; Sequence 1027, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1027
; LENGTH: 10317
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1027

Alignment Scores:

Pred. No.: 2,03e-05 Length: 10317
Score: 184.00 Matches: 155
Percent Similarity: 29.4% Conservative: 70
Best Local Similarity: 20.3% Mismatches: 228
Query Match: 4.6% Indels: 312
DB: 3 Gaps: 40

US-09-917-376-3 (1-740) x US-09-902-540-1027 (1-10317)

Qy 44 TrpAspAlaAlaAsnGlyArgTrpIlePro----- 53
Db 4722 TGGGATCGCAGCGGGAGGCTTTCATCATGACAGCGGACGGGCTCGGTCGCAAAACCAG 4663

```
QY 54 -----LeuLeuAspTrpValGly-----TrpAsn 61
Db 4662 CCCACGGGGTCTCTCAATGACACCGTGGCGCTCACCTCCGTCGGGACCGGGACCTGGGAG 4603
QY 62 AsnTrpGlyTyraAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysVal 81
Db 4602 GAGTGGCGCGTCCGGGACTCTCGCGCCCGCGCGCTCCGGCGCGGGTGGTGGCGCC 4543
QY 82 TrpAlaAlaValGly----- 86
Db 4542 TTCAGTATCGCGGACCGGCATCAGCGTGTCTTCGTGGAGAGAGCTGCACGGGTGCG 4483
QY 87 MetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAspGlnGly 106
Db 4482 AAGCTGTGGACACGTGGCATTCACTCAGCACT----- 4447
QY 107 AlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArgGly 126
Db 4446 CGGACCTGGCAG-----CAGTGGGGCGCGCGGGCCCTCCGCGCGT 4405
QY 127 MetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAlaPro 146
Db 4404 TCGGGCGCGGTGACGGCTCCCGGTCTCGGGCAGCATCTTGTCTCGAGGTACC 4345
QY 147 SerGlyLysGly-----LeuTrpArgSerThrAsp 156
Db 4344 TCGGGACGGCTCGCAGGGAGGCGCTGTCTGTCACCAACAGAGCTTTGGAGCCATGCCAG 4285
QY 157 SerGlyAlaThrTrpSerGlnMetThrAsnProAspValGlyThrTyrIleAlaAsn 176
Db 4284 GGGTGGCCCAAGTTGCTCCACCGGACGTGCGGACCGCGCGCTTCGCGCGGCC 4225
QY 177 ProThrAspThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAsp 196
Db 4224 CGCTGCACAGAG-----TCCGGGCTCATGTGATGATCGCGCGGT 4183
QY 197 LysSerSerSerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspPro 216
Db 4182 GCGAGCGCGCGCCATCGGC-----GCCGACCTG 4153
QY 217 AsnAsnProValPheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAla 236
Db 4152 GCGACGTGCTCCAGTACACCGCTCGACCAATGCC-----TGGGCGAGCGCGCGCTCGCTG 4096
QY 237 ProThr-----GlyPheIleProHisLysGlyValPheAspPro 249
Db 4095 CCGCGGTGCTCGCGGTGATGATGCTATCGCGCGCGCGGGATGTAT----- 4042
QY 250 ValAsnHisValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyrAspGlySerSer 269
Db 4041 -----CTCTTCGGAGGACGTCCAC-----GGTGGCGCGCGC 4009
QY 270 GlyAspValTrpLysPheSerValThrSerGlyThrTrpThrArgIle----- 285
Db 4008 AACGATCTGCTCGGTTCAGC-----CGGGGCGGTGGACACCGGTCCACCGCGCAGGGC 3955
QY 286 -----SerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyrSerGly 302
Db 3954 GCAGCGGGAGTCTCT-----CCGACGAT-----GGTGGG 3925
QY 303 LeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTrp 322
Db 3924 GTAGCTCAGAGCGAC-----CCACAGACA----- 3898
QY 323 ProAspThrIleIlePheArgSerThrAspGlyGlyAlaThrTrpThrArgIleTrpAsp 342
Db 3897 GCGACCTGCTCTCTCCACGGGAC-----GGGGCGGTGTGAC-----TTCGAC 3850
QY 343 TrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGluProTrp 362
Db 3849 GGAACGGGTGGCAGCGGGCGCACT-----TCGTCTCTCGCGCGGAGACC 3805
QY 363 LeuThrPheGlyValGlnProAsnProProValProSerProLysLeuGly----- 379
```

```
Db 3804 GTCACGCTGTGGGACGATTTCGGCGGAGTGACGAGCGGTTCAGTGGCGCACCATC 3745
QY 380 TrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThr 399
Db 3744 TGGGTG-----GTCGCGCCACCGCGCTTCAGCGGACACAGTCTTACGTC 3700
QY 400 GlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlyGlnIleHis 419
Db 3699 GGTCTGACG-----GGTGGG----- 3685
QY 420 IleAlaProMetValLysGlyLeuGluThrAlaValAsnAspLeuIleSerProPro 439
Db 3684 -----GTCGCCAGCTACAAGCTCAGCGGGTGGCG 3655
QY 440 SerGlyAlaProLeu-----IleSerAlaLeuGlyAspLeuGlyGly-----PheThrHisAla 457
Db 3654 CGGGCGTTCGCTCTCCGCTCTATGCTACTACGAGTCGGGAGGCTTTATTCTCAAG 3595
QY 458 AspValThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSer 477
Db 3594 GACGTGGGGAGGTGGGCGCGCTGTC-----CGGACACGACG 3556
QY 478 ValAspTyrAlaGluLeuAsnProSerIleValArgAlaGlySerPheAspProSer 497
Db 3555 TTGACATCGCG----- 3544
QY 498 SerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGln 517
Db 3544 ----- 3544
QY 518 GlySerGluProGlyGlyVal-----ThrThrGlyGlyThrValAlaAlaSerAla 534
Db 3543 ---CTGGAACCGCGCCCTCGGACACGACGACGACGACGACGACCTCGTCCCTCCCGC 3487
QY 535 AspGlySerArgPheValTrpAla----- 542
Db 3486 GATTGGCCCGCTTCATCAATCCCGCTATCGCGAGGGGTTCGCTCCAGTCAGGACTT 3427
QY 543 ProGlyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAla 562
Db 3426 CCGCGCGACCAATGGCCCGCCAGAGACGACGCGCGCGCACGTCGAGGACGTTCTCCGG 3367
QY 563 SerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThrPhe 582
Db 3366 ACCTAC---ATCCCTGCTCGTCCGCA-----CCCGCACCTCAG 3334
QY 583 TyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThrPheGlnPro 602
Db 3333 GCGCTCTCGTCGAGGCGACGCTC----- 3310
QY 603 ValAlaAlaGlyLeu-----ProSerSerGlyAla 612
Db 3309 GTGGGGCATCGCTCTGCGAGCGTTGACGCTGGGTCTACAATCTCCCGCGGTGCT 3250
QY 613 ValGlyValMetPheHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaSer 632
Db 3249 GTCGCGCGGTG-----GCGCTACCG----- 3229
QY 633 SerGlyLeuTyrHisSerThrAsnGlyGlySerTrpSerAlaIleThrGlyValSer 652
Db 3228 -----GATGGCATCCGGGCGGTG----- 3211
QY 653 SerAlaValAsnValGlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPhe 672
Db 3210 -----GCGCGGGGCTGAGTGAATGCGCGCACCGGG 3181
QY 673 ValValGlyThrIleGlyGlyValThrGlyAlaTyr----- 684
Db 3180 ACGGTGCGCGGTG-----GCGGAGGGGGCGTATCAGCTCAGCGCGCGCGCGGACG 3127
QY 685 -----ArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAspAsp 699
```


Qy	102	-----SerSerAspGlnGlyAlaThr	108
Db	3744581	TGGGCAATATTGGGTGGGCAATGCGGGCAGCAGCAACTAGCGGCTCGCAAACTGGGTG	3744582
Qy	109	TrpGlnIleThrProLeuProPheLysLeu-----Gly	119
Db	3744521	TGGGCAACATCGGTTTGGCAACACGGGTAGCAACAACATCGGGATCGGGTTGACCGGG	3744522
Qy	120	GlyAsnMetProGly-----ArgGlyMetGlyGluArgLeuAlaValAsp	134
Db	3744461	AC-AACCTGACTGGCATTTGGGGGCGCTGAATTACGAAACCGGTAAATCTCGGGTTGTTCAAC	3744462
Qy	135	ProAsnAsnAspAsnIleLeuThrPheGlyAlaProSerGlyLys-----GlyLeuThrPArg	153
Db	3744402	TCCGGCACCGGCAACATCGGTTCTCAATTTCGGGGACCGGCAACTTCGGGGTATTCAC	3744403
Qy	154	Ser-----ThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPheProAsp	169
Db	3744342	TCCGGCAGCTACAACACCGGTGTCGTTAAATCGGGGACGGCAGTACCGGGTTGTTCAAC	3744343
Qy	170	Val-----GlyThrTyrlle	174
Db	3744282	GTGTGGGTTCACACCGGTGTGGCCAACTGGGTAGCTATAACACGGGCAGCTTCAAC	3744283
Qy	175	AlaAsnProThrAspThrThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAla	194
Db	3744222	GGGGCAACACCAATACGGGTGGCTTCAACCCGGGCAACGTCACACCGCTGGTGAAC	3744223
Qy	195	PheAspLysSerSerSerSerLeu-----GlyGlnAlaSerLysThrIlePheVal	211
Db	3744162	ACCGGCAACACCAACACCGGCATCGCAACTCGGGCAATGTCAACACCGGGCGTTTCATC	3744163
Qy	212	GlyValAlaAspProAsnAsnProValPheTrpSerArgAsp-----	225
Db	3744102	TCG---GGCAATCTCAGCAACCGGTGCTGTGGCGGGTCACTACGAGGGCTGTGGGGG	3744103
Qy	226	GlyGlyAlaThrTrpGlnAlaValPro-----GlyAlaProThrGlyPhe	240
Db	3744045	CTCTCGGTGGATCGACCATTCGGGGCATCCCATTTGGTCTCGAGCTCAACGGCGGCGTC	3744046
Qy	241	IleProHisLysGlyValPheAspPro-----ValAsnHisValLeuTyrlleAlaThr	258
Db	3743985	GG-CCCCATCACCGGTGTCGGATCCAGATTTTGGCCACCATCCCGCTCAACATTACCA	3743986
Qy	259	SerAsnThrGlyGlyProTyrAspGlySerSerGlyAsp-----ValTrp	273
Db	3743926	AACCTTCAGCCTCGGCCC---GCTGGTCTGTCGGCAATCGTATGCCCGCTTTGGTGG	3743927
Qy	274	LysPheSerValThr-----	278
Db	3743869	CGGTACGGCCATACCTATCAGCGTCGGCCCCCATCACCATCTCGGCCCATCCCTGTTC	3743870
Qy	279	SerGlyThrTrpThrArgIleSerPro-----	287
Db	3743809	GGCTCAGA-ACITTCACACGACTTTCGCCGTCGSCCCCTCTTTGGCTTGGGGTCGTCA	3743810
Qy	288	-----ValProSerThrAspThrAlaAsn-----Asp	296
Db	3743750	ACATTTACAGGAATCGAAATCAAGATCTTCGGCGCAACGTCACCCCTCCCAATTAGTAACC	3743751
Qy	297	TyrPheGlyTyrlleSerGlyLeuThrIleAspArgGlnHisProAsnThr-----	312
Db	3743690	TTAATATCGACACACAGAAATTAACAGTCATTCGGGTGACGTCACATGGAGTACCCCGG	3743691
Qy	313	-----IleMetValAlaThrGlnIleSerTrp-----TrpPro	323
Db	3743630	CAGTAACGATCTCCCGAATGGCATCAGTATTCACCAATCCACTGCGCTGTGGCCA	3743631
Qy	324	AspThrIleIlePheArgSerThrAspGlyGlyAlaThrTrpThrArgIleTrpAspTrp	343
Db	3743570	-----GCGGTCGATCGGCACGCTGGGATTCA	3743571

```
QY 344 ThrSer-----TyrProAsnArgSerLeuArgTyrValLeuAspIleSerAla 359
Db 3743543 CGATCCCGGGCTTCCATTCCCGCTCGCGCTG-ACGATCGACATACACGGC 3743485
QY 360 Glu----- 360
Db 3743484 CAGATTGACGGCTTCAGACACCCCGCATCAGCATCGACCGCATCCCGTGAACCTCGC 3743425
QY 361 -----ProTrrPLeuThrPheGlyValGlnProAsnProProValPro 374
Db 3743424 GCCAGCGTCACTGTCGGCCCTATCTCTGATCAACCGCGTT-----AATATCCCGCGC-- 3743374
QY 375 SerProLysLeuGlyTrrMetAspGluAlaMetAlaIleAspPropheAsnSer---Asp 393
Db 3743373 ACCCGGGCTTTGGCAACACACACCGCTCCGTCGCGGTTTCTTCAACTCCGGCGAC 3743314
QY 394 ArgMetLeu-----TyrGlyThrGlyAlaThrLeuTyr-----Ala 405
Db 3743313 GGTGGGTGTCGGGCTTCGGGAATTCGTCGGGCGAGCTCGGGTTGTTGTAACACGCGC 3743254
QY 406 ThrAsnAspLeuThrLysTrrAspSerGly-----Gly 416
Db 3743253 CAGACCGAGGTGGTGGCGGGTTCGGTTTCGCCAATTTCCGTTTCGCTCGGATCGGT 3743194
QY 417 GlnIleHisIleAlaProMetValLysGlyLeuGluThrAlaValAsnAspLeuIle 436
Db 3743193 GTGCTGAATTCGCTCGGCTGTGTCGGGCTGTACACACCGCGCGG----- 3743146
QY 437 SerProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHis 456
Db 3743145 TTCCCGCGGGGACCCCGGGTGGTCTCGGATCGGCAATGTTGTGTAG----- 3743095
QY 457 AlaAspValThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThr 476
Db 3743094 -----CAGCTGTGCGGGTGTCTCGCGCGGACG 3743055
QY 477 SerValAspTyrAlaGluLeuAsnProSerIleValArgAlaGlySerPheAspPro 496
Db 3743064 GCA-----CTCAACGACAGCTCATCATCAATTCGCGTTGGCGGATGTG 3743020
QY 497 SerSerGlnProAsnAspArgHisValAlaPhe-SerThrAspGlyGlyLysAsnTrrPh 516
Db 3743019 GGCAGCGTA-----AAGCTCGGTTTCGGCAACGTCGGGACTTCAACCTCGGT 3742972
QY 516 eGln-----GlySerGluProGlyGlyValThrGln 527
Db 3742971 GCGCCCAATATCGCGACTTGAACGTGGTTTGGCAATCTCGCGCGCAACGTCTGG 3742912
QY 527 YGlyThrValAlaAla-----SerAlaAspGlySerArgPheValTrrAlaProGlyAs 545
Db 3742911 TTCGGCAATATCGCGATGCAACTTCGGGTTCGGCAATCGGGTCTGCGCGGGCGCTG 3742852
QY 545 pProGlyGlnProValTrrAlaValGlyPheGlyAsnSerTrrAlaAlaSer-----G1 564
Db 3742851 GCCCGGG-----TGGCAACATCGGGTTG 3742828
QY 564 nGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThrPheTrrAl 584
Db 3742827 GCGAATCGCGGACGCGCA-----AC 3742807
QY 584 aLeuSerAsnGlyThrPheTrrArgSerThrAspGlyGlyValThrPheGlnProValAl 604
Db 3742806 GTCGGCTTCGGCAACATGGGTGTGGCAACATCGGGTTCGGTAACCGCGCACCAACAC 3742747
QY 604 aAlaGlyLeu-----ProSerSe 610
Db 3742746 CTCGGATTGGGTGACCGGGGCAACACAGACTGGGATCGCGCTTGAACCTCGGT-GC 3742688
QY 610 rGlyAlaValGlyValMetPheHisAlaValProGlyLysGlyGlyAspLeuTrrPleuAl 630
Db 3742687 CGGCAACATCGGG-----TTGTCAACTCCGCCACCGCAACGTCGGG-----TTGTCAACTC 3742634
QY 630 aAlaSerSer-----GlyLeuTrrHisSer-----ThrAsnGlyGlySerSe 644
```

```
Db 3742633 CGGACCGGGAACTTCGGGTGTTTCAACTCGCGGACAGCTTCAACACCGCATCGGCAATGG 3742574
QY 644 rTrrSerAlaIleThrGlyValSerSerAlaValAsnValGlyPheGlyLysSerAlaPr 664
Db 3742573 CGGAACGGGAGTACTGGCTTTTCAATCGCGGTAAATTTCAATACCGGTGTCGCAACCC 3742514
QY 664 oGlySerSerTyrProAlaValPheValValGlyThrIleGlyGlyValThrGlyAlaTrr 684
Db 3742513 TGGGTTCGTACAAACCGGCGAGCTTCAATGTGGTGACACCAAC-----ACCGGTGGTTT 3742460
QY 684 rArgSerAspAspCysGlyThrTrrValLeuIleAsnAspAspGlnHisGlnTrrG1 704
Db 3742459 CAACCCGGGAGCATCAACACCGGCTGGTTCAACACCGGCAACGC-CAACACCGGCTCG 3742401
QY 704 yAsnTrrGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgArgValTrrIleGly 723
Db 3742400 CCAATTCGGGCA-----TGTGACACCGGCGCCCTCATGTGCGG 3742361

RESULT 9
US-09-328-352-1377
; Sequence 1377, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; NUMBER OF SEQ ID NOS: 8252
; LENGTH: 11679
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1377

Alignment Scores:
Pred. No.: 4,11e-05 Length: 11679
Score: 181.50 Matches: 204
Percent Similarity: 31.0% Conservative: 90
Best Local Similarity: 21.5% Mismatches: 344
Query Match: 4.5% Indels: 309
DB: 3 Gaps: 54

US-09-917-376-3 (1-740) x US-09-328-352-1377 (1-11679)
QY 7 ThrTrrSerAsnValAlaIleGlyGlyGlyPheValAspGly----- 21
Db 1627 ACGTGGACA---GTAAGTGTGCGGGTAGTGGTTTGGTTGCTGATGCAGATAAAACGATT 1683
QY 22 -----IleValPheAsnGluGlyAlaProGlyIleLeuTrrValArg---ThrAsp 37
Db 1684 GATGCTAAAGTAACGTTTACAGATGTCAGAGGTAATAGCAGCACTGTTAACGATACGCAA 1743
QY 38 IleGlyGlyMetTrrArgTrrAspAlaAlaAsnGlyArgTrrIleProLeuAspTrr 57
Db 1744 ATT-----TATACATTAGACACAGCT---GCTCTGCGCGCCAGTAAATGACCCA 1791
QY 58 ValGlyTrrAsnAsnTrrGlyTrrAsnGlyVal-----ValSerIleAlaAlaAspPro 75
Db 1792 GTT-----AACGGACAGACCAATACAGGTACAGCAACCT 1830
QY 76 IleAsnThrAsnLysValTrrAlaAlaValGlyMetTrrThrAsn----- 90
Db 1831 GGTTCACAGTAACAGTA-----ACCTATCTCTAATGTTGTCACACCAACA 1875
QY 91 -----SerTrrAsp---ProAsnAspGlyAlaIleLeuArgSer 102
Db 1876 GTTGTACAGACCGGACGCGAGTGTGTAGTACCAACCTCGGC-----CTTAATGAT 1929
QY 103 SerAspGlnGlyAlaThrTrrGlnIleThrProLeuProPheLysLeuGlyGlyAsn--- 121
```

1930 GGTGACGAAGTTGAGGCAATTGTTACAGATCCA-----GCAGGCAACCCA 1974
Db
122 ---MetProGlyArgGlyMetGlyGluArgLeuAlaValAlaAspProAsnAsn----- 137
Qy
1975 TCCTTGGCAGGTACAGCTACTGTTGAT-----GCAGTTGGTCCAAATACCGATGGTGT 2028
Db
138 -----AspAsnIleLeuTyrPheGlyAlaProSerGly 148
Qy
2029 AACTTTACGGTTGATTTCAGTAACAGCTGCAGATGATTAATGATCATGAGACGGTCAGGC 2088
Db
149 Lys -----GlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMet 164
Qy
2089 AACGTTACTGTTACTGTTGTTGATTAAGAAACGTTCCGGCAGATGAGCAAAATACAGTGTGTC 2148
Db
165 ThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGln 184
Qy
2149 ACTGTTGTAATCAATGGCCAGACGTATACCTGCAACTGTAGATAGCAGCAGGC----- 2202
Db
185 SerAspIleGlnGlyValTrpValAlaPheAspLysSerSerSerLeuGlyGln 204
Qy
2203 -----ACATGGACAGTAAGCGGTACCAGGTAGTGACTTAACCTCGCGAT 2244
Db
205 AlaSerLysThrIle----- 209
Qy
2245 GCAGTAAGACGATTGATGCTAAAGTAACGTTTACAGATGCGGCGAGGTAATAGCAGCAGT 2304
Db
209 ----- 209
Qy
2305 GTTAAACGATACACAAATATACATACATACATACATACATACATACATACATACATACAT 2364
Db
210 -----PheValGlyValAlaAspProAsnAsnProValPheTrp 222
Qy
2365 CCGGTTAAACGGGACAGACCCGATTACAGGTACGGCAGAGCCTGGTTCAACAGTAACCTGTG 2424
Db
223 SerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaPro---ThrGlyPheIle 241
Qy
2425 ACTTATCCAGATGCGCATGATACAAACAGCTGTTGTCAGACCGGATGCGCATGTCGACAGTA 2484
Db
242 ProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsnThr 261
Qy
2485 CCAACCAGGTTTAATGATGGC---GATAAAGTTACAGCAATGCTACAGATCCAGCA 2541
Db
262 GlyGlyProTyrAspGlySerSerGlyAspValTrpLysPheSerValThrSerGlyThr 281
Qy
2542 GGCAACCCA-----TCATTACCAGGTACA 2565
Db
282 TrpThrArgIleSerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyrSer 301
Qy
2566 GCTACTGTTGATGCACTGGTTCCTCAAAATACCGATGCTGTTAAT-----TTCACGGTTGAT 2619
Db
302 GlyLeuThrIleAspArg----- 307
Qy
2620 TCAGTAACGCTGACAATGATTAATGATCATCAGAAGCATCAGGCAACGTTACTGTTTACA 2679
Db
308 -----GlnHisProAsnThrIleMetValAlaThrGlnIle 319
Qy
2680 GGTGATTGAAACCGTTCCGGCAGATGCACAAATACAGTG---GTCACCTGTTGTGATC 2736
Db
320 SerTrpTrpProAspThrIlePheArgSerThrAspGlyGlyAlaThrTrpThrArg 339
Qy
2737 AATGGCCAGACGTTACTGCAACTGTAGATAGCAGCAGGC-----ACATGGACA--- 2787
Db
340 IleTrpAspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAla 359
Qy
2788 -----GTAAGCGTACCAGGTAGTGACTTGACTGCGGATGATGATGATGATGATGATGAT 2835
Db
360 GluProTrpLeuThrPhe----- 365
Qy
2836 GATGCCAAAGTAACGTTTACAGATGACGAGGTAAACAGCAGTGTAAACGATACACAC 2895
Db
366 -----GlyValGlnProAsnProProValProSerProLysLeuGlyTyr 380
Qy
2896 ACATATACAGTTGATACGTTGACCAAAATGCACCGGTG----- 2934
Db
381 MetAspGluAlaMetAlaIleAspProPheAsnSerAsp----- 393
Qy
2935 CTTGATCCGATCAATGCAACAGACCCAGGTGAGCGGTGAGCAGAGCCCTGGTTCAACAGTG 2994
Db
394 ArgMetLeuTyr-----GlyThrGlyAlaThrLeuTyrAlaThrAsnAsp----- 408
Qy
2995 ACTGTGACTTATCTCTGATGGCACTGTCACACAGTGTGTAGCAGGACCGGTAGTGTGG 3054
Db
409 -----LeuThrLysTrpAsp 413
Qy
3055 TCAGTACCAAAACCCAGGTAACTGGTGGATGGGATACAGTACTGCTGCAACAGCAACTGAC 3114
Db
414 SerGlyGlyGlnIleHisIle-----AlaPro 422
Qy
3115 CCTGCAGGCAACACTTCTATTGCCAGGTACAGGCACAGTTCCTCAGCAGACATCACAGCACCT 3174
Db
423 MetValLysGlyLeuGluGlnThrAlaValAsnAsp-----Leu 435
Qy
3175 GTGGTT---GCCTGGATGACGTTGACGAATGACAGCAGCAGCAGCAGTACCGGTAC 3231
Db
436 IleSerProSerGlyAlaProLeuIleSerAla----- 447
Qy
3232 GTGAACGATCCAGCAGCCACTGTAGTTGCAATGTGATGGCTGCTGACTATCCGGCAGTG 3291
Db
448 ---LeuGlyAspLeuGlyPheThrHisAlaAsp-----ValThrAla 461
Qy
3292 AACAAATGGTGAC---GGCACCTGGACGCTTGCAGACAATACACTTCTCGCTTAACCTGAT 3348
Db
462 ValProSerThrIlePheThrSerProValPheThrThrGly----- 475
Qy
3349 GGTCCACACACACTTACCTGCTGCTCAACAGATGCGAGCGCAATGCGAGGTACAGATACA 3408
Db
476 ---ThrSerValAspTyrAlaGluLeuAsnProSerIleIleValAlaGlySer 493
Qy
3409 GCAGTGGTGACGATTGATACCAAGCAATGCAATGCAATGCTGATTCGATCAATGCG 3468
Db
494 PheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLys 513
Qy
3469 ACTGACCCA-----GTGACGGTACAGCAGAGTGGTTCA 3504
Db
514 AsnTrpPheGlnGlySerGluProGlyValThrThrGlyGlyThrValAlaAlaSer 533
Qy
3505 ACC---GTGACTGTGACTTATCCTGATGCAACCACT---GCAACAGTGGTAGCAGGC 3555
Db
534 AlaAspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProVal----- 550
Qy
3556 ACAGATGGTAGC-----TGTCAGTACCAACCCAGGTAACTCTGGTAGATGGTGTAT 3606
Db
551 ---ValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsn 569
Qy
3607 ACAGTAACCTGCAACA-----GCAACTGACCTGCGAGCAATACATCATTCGCCAGGTACA 3660
Db
570 AlaGlnIleArgSerAspArgValAsnProLysThrPheTyrAlaLeuSerAsnGlyThr 589
Qy
3661 GGCACAGTCTTCGACAGATACAGACACCT-----GTGGTTGGCTGGATCGCTGTG 3714
Db
590 PheTyrArgSerThrAspGly-----GlyValThrPheGlnProValAlaAlaGlyLeu 607
Qy
3715 ACCAATGACAGCAGCAGCAGCAGTTCACGGGTACAGTGAACGATCCGACAGCAGCTGTAGTG 3774
Db
608 ProSerSerGlyValAlaValGlyValMetPheHisAlaValProGlyLysGluGlyAspLeu 627
Qy
3775 GTGAAT-----GTAGATGGCAGTCACTGCTCGGCGAGTGAACAAATGCGCAGCGC---ACC 3825
Db
628 TrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsnGlyGlySerSerTrpSerAla 647
Qy
3826 TGGACGCTTGCAGACATACACTTCCAGTGTGGCAGAGCTCCACACACC---ATTACC 3882
Db
648 IleThrGlyValSerSerAlaValAsnValGly-----PheGly 660
Qy
3883 GTGACTGCAACAGATGCAGCAGGCAATGCGAGGTACAGATACAGCAGTGGTGACGATTGAT 3942
Db

```
QY 661 LysSerAlaProGlySerSerTyrProAlaValPheValValGlyThrIleGlyGlyVal 680
Db 3943 ACCACAGCACCAATGCA-----CCGCTACTTGATCCAATCAATGAGCTGACCCAGTG 3996
QY 681 ThrGlyAlaTyrArgSerAspCys-----GlyThrThrTyr 693
Db 3997 AGCGGTACACAGAGAGTGGTTCAACCGTGACTGTGACTTATCTGTATGCGACCACTGCA 4056
QY 694 ValLeuIle-----AsnAspAspGlnHisGlnTyrGlyAsnTyrGlyGlnAlaIleThr 711
Db 4057 ACAGTGTAGTAGCAGCAGACAGATGGTAGTGGTGCAGTACCAACCAGGTAACTGGTAGAT 4116
QY 712 GlyAspHisAlaAsnLeuArgArgValTyrIleGlyThrAsn-----GlyArg 727
Db 4117 GGTGATACAGTAAGTCAACAGCAACTGACCCCTGCAGGCAATACATCATTCGCCAGGTACA 4176
QY 728 GlyIleValTyrGlyAspIle 734
Db 4177 GGCACAGTCTCTGCAGACATC 4197

RESULT 10
US-09-949-016-11814
; Sequence 11814, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11814
; LENGTH: 29927
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(29927)
; OTHER INFORMATION: n = A, T, C or G
US-09-949-016-11814

Alignment Scores:
Pred. No.: 0.000436 Length: 29927
Score: 177.00 Matches: 189
Percent Similarity: 28.8% Conservative: 56
Best Local Similarity: 22.2% Mismatches: 292
Query Match: 4.4% Indels: 314
DB: 3 Gaps: 48

US-09-917-376-3 (1-740) x US-09-949-016-11814 (1-29927)
QY 14 GlyGlyGlyGlyPheValAspGly-----IleValPheAsnGluGly 27
Db 19367 GGGGGCTCTGGCAGGGGGGACGGCTGCTTCTTCCCGCTTTTATTCACAGGGGAC 19426
QY 28 AlaProGlyIleLeuTyrValArgThr-----AspIleGlyGlyMet 41
Db 19427 AGCGCTGGGATTGTATTGTGGCGCGGTGTTTGGCTGAGGGTGCAGGGACTT-GGGGGG- 19482
QY 42 TyrArgTyrAspAlaAlaAsn----- 48
Db 19483 TGGCGGTGGGGAGCGGAGGAGGTATAAAGTATAAATCATAAGTAAACAACTCAGAAATG 19542
QY 49 -----GlyArgTyrIleProLeu-LeuAspTyrTrpValGlyTyrAsnAsnTyrGlyTyrAs 66
Db 499 -----GlyArgTyrIleProLeu-LeuAspTyrTrpValGlyTyrAsnAsnTyrGlyTyrAs 66
```

```
Db 19543 GACCCCGAGCGCTGGTCCCGCTAGCTCTCCAGCTCTCCCTGGCCGACCCCGAAGGAGA 19602
QY 66 nGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTyrAlaAlaValG1 86
Db 19603 GGGGTCCGCACTCCCTCCGCGGTCTCTCTCTC----- 19633
QY 86 yMetTyrThrAsnSerTyrAspProAsnAspGlyAla-----IleLeuArgSerSe 103
Db 19634 -----CTGGGTACTCTGGCCCTTGAGGTGGGGAAACGAGCCTACTTCTTGTACCGTCTTT 19685
QY 103 rAspGlnGlyAlaThr-----TyrGlnI1 111
Db 19687 TGGCGACGCGCGGACCCAGTGAATAGCCGCTTGGACCCGCGACGCTGCTGGCTTTG 19746
QY 111 eThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArgGlyMetGly----- 128
Db 19747 GCGACCGGAGTCTTGG-----GGACCTGGTGTCCCGGAGAGGCTTGGGTACTCTGTTCTCT 19800
QY 129 -----GluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyr----- 142
Db 19801 TCCCGCGGAGAGGCTTGGGACCTGCTGCTCCGCGGAGAGGCTTGGGTACTCTGTTCTCT 19860
QY 143 -----PheGlyAlaProSerGlyLysGlyLeu----- 151
Db 19861 GGAAGAGCTTGGACACCTGCTGCTGGGAGGCTTGGGACCTGCTGCTCTCTGGAGA 19920
QY 152 -----TyrArgSerThrAspSerGlyAlaThrTyrSerGlnMetThrAsnPheProAspVa 170
Db 19921 GGCTTGGAGATCTGTTCTCTGGGAGGCTTGGGA----- 19957
QY 170 lGlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGlnSerAspIleGlnGlyVa 190
Db 19958 -----CCTGGTGTCTCTGGAGA 19974
QY 190 lValTyrValAlaPheAspLysSerSerSerSerSerSerSerSerSerSerSerSerSer 210
Db 19975 GGCTTGG-----GGACCTGGTGCACCTTGGAGAGGCTTGGAGACCTGGTGT 20019
QY 210 eValGlyValAlaAspProAsnAsnPro-ValPheTyrSerArg----- 224
Db 20020 TCTGGGAGAGGCT-----TGGGACCTGGTGTCTCTGGAGAGGCTTGGGAGACCTGGTGT 20073
QY 225 -----AspGlyGlyAlaThrTrp-----GlnAlaValProG 235
Db 20074 CTCTGGAAGAGGCTTGGACACCTGCTGACCCCGGAGGCGCTTGGGATCTGCTGCCGG 20133
QY 235 lAlaProThrGlyPheIleProHisLysGlyValPheAspProValAsnHisValLeuT 255
Db 20134 GAGAGCCTTGGGA-----CCTGGTGTCTCTG----- 20159
QY 255 yrlleAlaThrSerAsnThrGlyGlyProTyrAspGlySerSerSerSerSerSerSerSer 275
Db 20160 -----GGAGAGGCTTGGGAGCCT-----GGTGACCTTGGAGAGGCTTGGGAGC 20202
QY 275 heSerValThrSerGlyThrTyrThrArgIleSerProValProValProValProValPro 295
Db 20203 CTGGTGTCTCTGAGAGACCTTGGGA-----TCTGGTGTCTCCAGAGAGGCTTGGGAGC 20256
QY 295 snAspTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetV 315
Db 20257 CT-----GGTGTCTCTGGAAGAGGCTTGGAC----- 20282
QY 315 alAlaThrGlnIleSerTyrTrpPro-----AspThrIleIlePheArgS 330
Db 20283 -----ACCTGGTGTCTCTGGGAGAGGCTTGGGAGCCTGGTG----- 20318
QY 330 erThrAspGlyGlyAlaThrThrThrArgIleTyrAsp-TyrThrSerTyrProAsnArg 349
Db 20319 -----TCTGGGAGAGGCTTGGGAGCCTGGTGT-----CCTGGGAGA 20355
QY 350 SerLeu-ArgTyrValLeuAspIleSerAlaGluProTyrPheThrPheGlyValGlnPr 369
Db 20356 GGCTTGGAGA-----TCTGGTGTGAGCGGAGAGGCTTGGGA----- 20392
```

```
QY 369 oAsnProValProSerProLysLeuGlyTrpMetAspGluAlaMetAlaIleAspPr 389
Db 20393 ---CCTGGTCTCCGGGAGGCTTGGGTCTCCGGAGAGCTTGGACACC 20448
QY 389 o-PheAsnSerAspArgMetLeuTyrglyThrGlyAlaThrLeuTyraThrAsnAspL 409
Db 20449 TGGTCTCCAGAGAGGCTT---GGGACCTGGTGTACCTTGGAGAGGCTGGGGACC 20502
QY 409 euThrLysTrpAspSerGlyGlnIleHisIleAlaProMetValLysGlyLeuGluG 429
Db 20503 TGGTGACCCGGGAGAGCCTTGGG----- 20525
QY 429 luThrAlaValAsnAspLeuIleSer-----ProProSerGlyAlaProLeuIleSerA 447
Db 20526 ---GACCTGGTGTCTGGGAGAGCCTTGGGGACCT----- 20558
QY 447 laLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThrAlaValProSerThrIleP 467
Db 20559 ---GGTGACCTTGGAGAGCTTGGGACCTGGTGTCTCGGAGTGCCTTGGGACCT- 20612
QY 467 heThrSerProValPheThrThrGlyThrSerValAspTyraAlaGluLeuAsnProSerI 487
Db 20613 ---AGTGACCCGGGAGAGCCTTGGGACCTGTGTCTCCGGGAGAGCCTTGGGACCTGGT 20670
QY 487 leIleValArgAlaGlySerPheAspProSerSerGlnProAsnAspArgHisValAlaP 507
Db 20671 TCCTG-----GGAGACCTTGGGGATCTGGTGTCTCTG- 20702
QY 507 heSerThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGlyValThrG 527
Db 20703 ---GGGAGAGCTGGGGGACCTGTGTCTCGGAGAGAGCCTTGGGGACCTGGTGACC 20757
QY 527 lyGlyThrVal---AlaAlaSerAlaAspGlySerArgPheValTrpAlaProGlyAspP 546
Db 20758 CGGGAGGCTTGGACACCTGGTGTCCGGGAGAG-----GCTTGGGACCTGGTGACC 20811
QY 546 roGlyGln-----ProValValTyraAlaValGlyPheGlyAsnSerTrpAlaAla 563
Db 20812 CGGGAGACCTTGGGACCTGGTGTCTCGGGAGAGGCTGGGG---ACCTGGTGTCTCG 20868
QY 563 erGln-----GlyValProAlaAsnAlaGlnIleArgSerAspArgValA 578
Db 20869 GGAGAGACCTTGGGACCTGGTGTACCCGGGAGAGCTTGG----- 20909
QY 578 snProLysThrPheTyraAlaLeuSerAsnGlyThrPheTyraArgSerThrAspGly---- 596
Db 20910 -----ACACCTGGTGTGCC 20922
QY 597 --GlyValThrPheGlnProValAlaAlaGlyLeuPro---SerSerGlyAlaValGlyV 615
Db 20923 CGGGAGGCTTGGGAGCTGGTGTCCGGGAGAGCCTTGGGGACAGTACCTTGGAG 20982
QY 615 alMetPheHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyL 635
Db 20983 AGGCTTGGGGA-----CCTGGTGTCTTGGAGAGCCTTGGGACCTGGTGTCTCGGA- 21035
QY 635 euTyraHisSerThrAsnGlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlaV 655
Db 21036 ---GAGGTTACGGGGGCTGTGTGGGGAGAGACGTTGTGAGGCAAAAGTCCC 21084
QY 655 alAsn----- 656
Db 21085 TGAATCCCTCGAAGAGAGCCATCGGGAGCTCCCTCCCTGAGGGCGGTCCATTTGTGGACC 21144
QY 657 -----ValGlyPheGlyLysSerAlaProGlyLysSerSerTyraProAlaV 671
Db 21145 CCCCTCCATCGGCTTTCAGAGGAGCTGTTCGGATTCCCTTGGCCCGGCT---CCCGGG 21201
QY 671 alPheValValGlyThrIleGlyValThrGlyAlaTyraArgSerAspCysGlyT 691
Db 21202 ATGCATCCAGTGGCAGCGCCCAATTCGTGGCCAGGGGAGAGGAAAGCGGGTGTGGG 21261
```

```
QY 691 hrThrTrpValLeuIleAsnAspAspGlnHisGlnTyrglyAsnTrpGlyGlnAlaIleT 711
Db 21262 TGGTC-----TCCACGG 21273
QY 711 hrGlyAspHisAlaAsnLeuArgValTyraIleGlyThrAsnGlyArgGlyIleValT 731
Db 21274 CTGGAGAGGGGGGACGCTCCCTAGG-----GGAGAGAGGACCGTTGGGGGTTTC 21324
QY 731 yrGlyAspIleGlyGlyAla 737
Db 21325 CGGGGGCGGGGGCGGAGCA 21344

RESULT 11
US-09-949-016-17474
; Sequence 17474, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17474
; LENGTH: 29927
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(29927)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17474
```

Alignment Scores:
Pred. No.: 0.000436 Length: 29927
Score: 177.00 Matches: 189
Percent Similarity: 28.8% Conservative: 56
Best Local Similarity: 22.2% Mismatches: 292
Query Match: 4.4% Indels: 314
DB: 3 Gaps: 48

US-09-917-376-3 (1-740) x US-09-949-016-17474 (1-29927)

```
QY 14 GlyGlyGlyGlyPheValaspGly-----IleValPheAsnGluGly 27
Db 19367 GGGGGCTCTGGCAGGCGGAGCGGCTCCTCTTCCACCGTTTTTATTTCAAGGGAC 19426
QY 28 AlaProGlyIleLeuTyraValargThr-----AspIleGlyGlyMet 41
Db 19427 AGGCTGGGGAITGTATTGTGGCGCGTGTGGCTGAGGTGACGAGCTT-GGGGG- 19482
QY 42 TyraArgTrpAspAlaAlaAsn----- 48
Db 19483 TGGCGGTGGGAGCGCGGAGGTATAAAGCTATATAACTAATCACTAAGTAACAACCTCAGAAATG 19542
QY 49 -----GlyArgTrpIleProLeu-LeuAspTrpValGlyTrpAsnAsnTrpGlyTyra 66
Db 19543 GACCCCGAGCGCTGGTGGCGGTAGCTCTCCAGCTCTCCCTGGCCCGGCGGAGGAGA 19602
QY 66 nGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaValG 86
Db 19603 GGGGTCCGCATCCCTCCCGGGTCTCTCTC----- 19633
QY 86 yMetTyraThrAsnSerTrpAspProAsnAspGlyAla-----IleLeuArgSerSe 103
Db 86 yMetTyraThrAsnSerTrpAspProAsnAspGlyAla-----IleLeuArgSerSe 103
```

Db 19634 -----CTGGGTACCTGGCTTGGAGTGGGGAGACGAGCCTACTTCTTGTACCGTCTTT 19686
 Qy 103 rAspGlnGlyAlaThr-----TrpGlnI 111
 Db 19687 TGCGAGCGGGAGCCAGTGAATTAGCGCTTGGAGCCGCGAGGCTCGCTGGCTTG 19746
 Qy 111 eThrProLeuProPheIysLeuGlyGlyAsnMetProGlyArgGlyMetGly-----128
 Db 19747 CGCACCAGACTCTGG-----GGACCTGTGTCCCGGAGAAACCTTGGGACCTGGTA 19800
 Qy 129 -----GluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyr-----142
 Db 19801 TCCCGGGAGAGCTTGGGACCTGGTGTCCCGGAGAGGCTTGGGTACTGTGTCTCT 19860
 Qy 143 -----PheGlyAlaProSerGlyLysGlyLeu-----151
 Db 19861 GGAAGAGGCTTGACACCTGGTGTCTCGGAGAGGCTTGGACCTGTGTCTCGGAGA 19920
 Qy 152 -----TrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPheProaspVa 170
 Db 19921 GGCTTGGAGATCTGTGTCTCTGGGAGAGGCTTGGGGA-----19957
 Qy 170 lGlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGlnSerAspIleGlnGlyVa 190
 Db 19958 -----CCTGGTGTCTCTGGAGA 19974
 Qy 190 lValTrpValAlaPheAspIysSerSerSerLeuGlyGlnAlaSerIysThrIlePh 210
 Db 19975 GGCTTGG-----GGACCTGTGTACCTTGGAGAGGCTTGGAGACCTGTGTGT 20019
 Qy 210 eValGlyValAlaAspProAsnAsnPro-ValPheTrpSerArg-----224
 Db 20020 TCTGGGAGAGGCT-----TGGGAGACCTGTGTCTGGGAGAGGCTTGGGACCTGTGT 20073
 Qy 225 -----AspGlyGlyAlaThrTrp-----GlnAlaValProG 235
 Db 20074 CTCTGAAGAGGCTTGACACCTGGTGTACCCGGAGGCGCTTGGGGATCTGGTGTCCCG 20133
 Qy 235 lAlaProThrGlyPheIleProHisIysGlyValPheAspProValAsnHisValLeut 255
 Db 20134 GAGAGCCTTGGGA-----CCTGGTGTCTGT-----20159
 Qy 255 yrIleAlaThrSerAsnThrGlyGlyProTrpAspGlySerSerGlyAspValTrpIysP 275
 Db 20160 -----GGAGAGGCTTGGGACCT-----GGTGACCTTGGAGAGGCTTGGGAC 20202
 Qy 275 heSerValThrSerGlyThrTrpThrArgIleSerProValProSerThrAspThrAlaA 295
 Db 20203 CTGGTGTCTTGAGAGCCTTGGGGA-----TCTGTGTCTCCAGAGAGGCTTGGGGAC 20256
 Qy 295 snAspTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetV 315
 Db 20257 CT-----GGTGTCTTGGAGAGGCTTGGAC-----20282
 Qy 315 alAlaThrGlnIleSerTrpTrpPro-----AspThrIleIlePheArgs 330
 Db 20283 -----ACCTGTGTCTCGGAGAGGCTTGGGACCTGGTG-----20318
 Qy 330 erThrAspGlyGlyAlaThrTrpThrArgIleTrpAsp-TrpThrSerTyrProAsnArg 349
 Db 20319 -----TCTGGGAGAGGCTTGGGACCTGGTGT-----CCTGGGAGA 20355
 Qy 350 SerLeu-ArgTyrValLeuAspIleSerAlaGluProTrpLeuThrPheGlyValGlnPr 369
 Db 20356 GGCCTTGGGA-----TCTGTGTAGCCGGAGAGGCTTGGGGA-----20392
 Qy 369 oAsnProValProSerProIysLeuGlyTyrMetAspGluAlaMetAlaIleAspPr 389
 Db 20393 -----CCTGTGTCTCGGAGAGGCTTGGGACCTGTGTCTCGGAGAGGCTTGGACACC 20448
 Qy 389 o-PheAsnSerAspArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspL 409
 Db 20449 TGTGTCTCCAGAGAGGCTT-----GGGAGCCTGTGTGACCTTGGAGAGGCTTGGGAGC 20502

Qy 409 euThrLysTrpAspSerGlyGlyGlnIleHisIleAlaProMetValIysGlyLeuGluG 429
 Db 20503 TGGTACCCGGAGAGCCTTGGG-----20525
 Qy 429 luThrAlaValAsnAspLeuIleSer-----ProProSerGlyAlaProLeuIleSerA 447
 Db 20526 -----GACTGTGTCTCTGGGAGAGCCTTGGGACCT-----20558
 Qy 447 laLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThrAlaValProSerThrIleP 467
 Db 20559 -----GGTGACCTTGGAGAGCCTTGGGACCTGTGTCTCGGAGTGTCTTGGGACCT- 20612
 Qy 467 heThrSerProValPheThrGlyThrSerValAspTyrAlaGluLeuAsnProSerI 487
 Db 20613 --ACTGACCCGGAGAGGCTTGGGACCTGTGTCTCCGGAGAGGCTTGGGACCTGGTG 20670
 Qy 487 leIleValArgAlaGlySerPheAspProSerSerGlnProAsnAspArgHisValAlaP 507
 Db 20671 TCCTG-----GGAGAGCCTTGGGAGTCTGTGTCTCTG- 20702
 Qy 507 heSerThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGlyValThrThrG 527
 Db 20703 -----GGGAGAGCCTTGGGAGCCTGTGTCTCGGAGAGAGCCTTGGGAGCCTGGTGACC 20757
 Qy 527 lyGlyThrVal---AlaAlaSerAlaAspGlySerArgPheValTrpAlaProGlyAspP 546
 Db 20758 CGGAGAGGCTTGGACACCTGTGTCTCCGGAGAG-----CCTTGGGAGCCTGGTGACC 20811
 Qy 546 roGlyGln-----ProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlas 563
 Db 20812 CGGAGAGCCTTGGGAGCCTGTGTCTCTCGGAGAGGCTGGGG--ACCTGTGTCTCG 20868
 Qy 563 erGln-----GlyValProAlaAsnAlaGlnIleArgSerAspArgValA 578
 Db 20869 GGAGAGAGCCTTGGGAGCCTGTGTCTCGGAGAGAGCTTGG-----20909
 Qy 578 snProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGly---- 596
 Db 20910 -----ACACTGGTGTCTC 20922
 Qy 597 --GlyValThrPheGlnProValAlaAlaGlyLeuPro---SerSerGlyAlaValGlyV 615
 Db 20923 CGGAGAGGCTTGGGAGCCTGTGTCTCCGGAGAGCCTTGGGACAGGTGACCTGGAG 20982
 Qy 615 alMetPheHisAlaValProGlyLysGlyGlyAspLeuTrpLeuAlaAlaSerSerGlyL 635
 Db 20983 AGGCTTGGGA-----CCTGTGTCTCTTGGAGAGGCTTGGGACCTGTGTCTCGGGA- 21035
 Qy 635 euTyrHisSerThrAsnGlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlaV 655
 Db 21036 -----GAGTTACGGGGCTGTGTGGGAGAGAACGTTGTGAGCCAAAGTCCC 21084
 Qy 655 alAsn-----656
 Db 21085 TGAATCCTCGAAGAGAGCGCATCGGAGCTCCCTCGAGGCGTTCATTTGTGGACC 21144
 Qy 657 -----ValGlyPheGlyLysSerAlaProGlySerSerTyrProAlaV 671
 Db 21145 CCCCTCCATCGCTTTTGCAGGAGCTGTTCGGATTCCCTGGCCCGCT---CCCGGG 21201
 Qy 671 alPheValValGlyThrIleGlyValThrGlyAlaTyrArgSerAspAspCysGlyT 691
 Db 21202 ATGCATCCAGTGGCAGCGCCAAATTCCTGGCCAGGGGAGAGGAAAGCGCGGTGTGGG 21261
 Qy 691 hrThrTrpValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleT 711
 Db 21262 TGTCTC-----TCCACGG 21273
 Qy 711 hrGlyAspHisAlaAsnLeuArgValTyrIleGlyThrAsnGlyArgGlyIleValT 731
 Db 21274 CTGGAGAAGGGCGACGCTCCCTAGG-----GGAGAAGAGGACGACGTTGGGGGTTTC 21324

QY 731 yrGlyAspIleGlyGlyAla 737
 Db 21325 CGGGGGCGGGGGCGGAGCA 21344

RESULT 12
 US-09-949-016-17475
 ; Sequence 17475, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17475
 ; LENGTH: 29927
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(29927)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-949-016-17475

Alignment Scores:
 Pred. No.: 0.000436 Length: 29927
 Score: 177.00 Matches: 189
 Percent Similarity: 28.8% Conservative: 56
 Best Local Similarity: 22.2% Mismatches: 292
 Query Match: 4.4% Indels: 314
 DB: 3 Gaps: 48

US-09-917-376-3 (1-740) x US-09-949-016-17475 (1-29927)

QY 14 GlyGlyGlyGlyPheValAspGly-----IleValPheAsnGluGly 27
 Db 19367 GGGGGCTCGGAGGGCGGAGCGGTGGCTCCCTCTTCACCGTTTATTCAAGGGGAC 19426

QY 28 AlaProGlyIleLeuTyrValArgThr-----AspIleGlyGlyMet 41
 Db 19427 AGGCTGGGGAATTGATTATTTGGGCGCGTGTGGCTGAGGCTGCAGGACTT-GGGGGG-- 19482

QY 42 TyrArgTTrpAspAlaAlaAsn----- 48
 Db 19483 TGGCGGTGGGAGCGCGGAGGTATAACGTATAATCAATAAGTAAACAACTCAGAAATG 19542

QY 49 -----GlyArgTTrpIleProLeu-LeuAspTrpValGlyTTrpAsnAsnTTrpGlyTyrAs 66
 Db 19543 GACCCGAGCGCTGTCCCGCTAGCTCTCCAGCTCTCCCTGGCCAGCGCCGGAAGGAGA 19602

QY 66 nGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTTrpAlaAlaValG1 86
 Db 19603 GGGGTCCGCATCCCTCCCGGTTCTCTCTC----- 19633

QY 86 yMetTyrThrAsnSerTTrpAspProAsnAspGlyAla-----IleLeuArgSerSe 103
 Db 19634 -----CTGGGTACCTGGCTTGAAGTGGGGAGACGAGCTTACTTCTGTACCTCTTT 19686

QY 103 rAspGlnGlyAlaThr-----TTrpGlnI1 111
 Db 19687 TGCCGACGGCGGAGCCAGTGAATTAGCGCTTGGACCGCAGCGCTGCTGCTGTTG 19746

QY 111 eThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArgGlyMetGly----- 128
 Db 20525

19747 CGCACCGGAGTCTTGG-----GGACCTGGTGTCCCGGAAAAAAGCTTGGGACCTGGTA 19800
 QY 129 -----GluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyr----- 142
 Db 19801 TCCCGGGAGAGGCTTGGGACCTGGTGTCCCGGAGAGCTTGGGTACCTGGTTTCTCT 19860
 QY 143 -----PheGlyAlaProSerGlyLysGlyLeu----- 151
 Db 19861 GGAAGAGGCTTGGACACCTGGTGTCTGGAGGGCTTGGGACCTGGTGTCTCTGGGAGA 19920
 QY 152 -----TTrpArgSerThrAspSerGlyAlaThrTTrpSerGlnMetThrAsnPheProAspVa 170
 Db 19921 GCCTTGGAGATCTGTGTCTCTGGAGAGGCTTGGGA----- 19957
 QY 170 lGlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGlnSerAspIleGlnGlyVa 190
 Db 19958 -----CCTGGTGTCTCTGGAG 19974
 QY 190 lValTTrpValAlaPheAspLysSerSerSerSerLeuGlyGlnAlaSerLysThrIlePh 210
 Db 19975 GCCTTGG-----GGACCTGGTGTGACCTTGGAGAGCTTGGAGAGCTGGTGT 20019
 QY 210 eValGlyValAlaAspProAsnAsnPro-ValPheTTrpSerArg----- 224
 Db 20020 TCTGGGAGAGGCT-----TGGGACCTGGTGTCTGGAGAGGCTTGGGACCTGGTGT 20073
 QY 225 -----AspGlyGlyAlaThrTTrp-----GlnAlaValProG 235
 Db 20074 CTCTGGAGAGGCTTGGACACCTGGTGTGACCGGGAGGCTTGGGATCTGGTGTCCCG 20133
 QY 235 lyAlaProThrGlyPheIleProHisLysGlyValPheAspProValAsnHisValLeuT 255
 Db 20134 GAGAGCTTGGGA-----CCTGGTGTCTCTG 20159
 QY 255 yrIleAlaThrSerAsnThrGlyGlyProTyrAspGlySerSerGlyAspValTTrpLysP 275
 Db 20160 -----GGAGAGGCTTGGGACCT-----GGTACCTTGGAGAGCTTGGGAC 20202
 QY 275 heSerValThrSerGlyTTrpThrArgIleSerProValProSerThrAspThrAlaA 295
 Db 20203 CTGGTGTCTCTGAGAGAGCTTGGGA-----TCTGGTGTCCAGAGAGGCTTGGGAC 20256
 QY 295 snAspTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetV 315
 Db 20257 CT-----GGTGTCTCTGAGAGGCTTGGAC----- 20282
 QY 315 alAlaThrGlnIleSerTTrpPro-----AspThrIleIlePheArgS 330
 Db 20283 -----ACCTGGTGTCTCTGGGAGAGGCTTGGGACCTGGTG----- 20318
 QY 330 exThrAspGlyGlyAlaThrTTrpThrArgIleTTrpAsp-TTrpThrSerTyrProAsnArg 349
 Db 20319 -----TCTGGGAGAGGCTTGGGACCTGGTGT-----CCTGGGAGA 20355
 QY 350 SerLeu-ArgTyrValLeuAspIleSerAlaGluProTTrpLeuThrPheGlyValGlnPr 369
 Db 20356 GCCTTGGAGA-----TCTGGTGTAGCCGGGAGAGGCTTGGGA----- 20392
 QY 369 oAsnProProValProSerProLysLeuGlyTTrpMetAspGluAlaMetAlaIleAspPr 389
 Db 20393 -----CCTGGTGTCCCGGAGAGGCTTGGGACCTTGGTGTCCCGGAGAGGCTTGGACAC 20448
 QY 389 o-PheAsnSerAspArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspL 409
 Db 20449 TGGTGTCCCGAGAGAGGCTT-----GGGACCTGGTGTGACCTTGGAGAGGCTTGGGAC 20502
 QY 409 euThrLysTTrpAspSerGlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluG 429
 Db 20503 TGGTGTAGCCGGGAGAGGCTTGGG----- 20525
 QY 429 luThrAlaValAsnAspLeuIleSer-----ProProSerGlyAlaProLeuIleSerA 447
 Db 20526 -----GACCTGGTGTCTCTGGGAGAGGCTTGGGACCT----- 20558

```

Qy 447 laLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThrAlaValProSerThrIleP 467
Db 20559 -----GGTGACCTTGGAGAGCTTGGGAGACCTGTGTCTCGGAGTCCCTTGGGACCT- 20612
Qy 467 heThrSerProValPheThrThrThrSerValAspTyrAlaGluLeuAsnProSerI 487
Db 20613 ---AGTGACCGGAGAGCTTGGGAGACCTGTGTCTCGGAGAGCTTGGGAGCTCGTG 20670
Qy 487 lelleValArgAlaGlySerPheAspProSerGlnProAsnAspArgHisValAlaP 507
Db 20671 TCCTG-----GGAGAGCTTGGGATCTGGTCTCTG- 20702
Qy 507 heSerThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGlyValThrThrG 527
Db 20703 ---GGGAGAGCTTGGGAGACCTGTGTCTCGGAGAGCTTGGGAGCTTGGGAGCTG 20757
Qy 527 lyGlyThrVal---AlaAlaSerAlaAspGlySerArgPheValTrpAlaProGlyAspP 546
Db 20758 CGGAGAGCTTGGACACCTGGTGTCTCGGAGAG-----GCTTGGGAGCTTGGTGACC 20811
Qy 546 roGlyGln-----ProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaS 563
Db 20812 CGGAGAGCTTGGGAGCTTGGTGTCTCGGAGAGCTGCGG-----ACCTGTGTCTCG 20868
Qy 563 erGln-----GlyValProAlaAsnAlaGlnIleArgSerAspArgValA 578
Db 20869 GGAGAGAGCTTGGGAGCTTGGTGTCTCGGAGAGCTTGG----- 20909
Qy 578 snProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGly---- 596
Db 20910 -----ACACCTGGGTGTC 20922
Qy 597 --GlyValThrPheGlnProValAlaAlaGlyLeuPro---SerSerGlyAlaValGlyV 615
Db 20923 CGGAGAGCTTGGGAGCTTGGTGTCTCGGAGAGCTTGGGAGACCGAGGTGACCTTGGAG 20982
Qy 615 alMetPheHisAlaValProGlyLysGlyAspLeuTrpLeuAlaAlaSerSerGlyL 635
Db 20983 AGGCTTGGGA-----CCTGTGTATCTTGAGAGGCTTGGGAGCTTGGGAGCTTGGGA- 21035
Qy 635 euTyrHisSerThrAsnGlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlav 655
Db 21036 ---GAGTTACGGGGCTGTGTCTGGGAGAGAACCTTGTGAGCCAAAGTCCC 21084
Qy 655 alaSn----- 656
Db 21085 TGAATCCTCGAGAGAGCGCATCGGAGCTCCCTCGAGGCGTTCATTTGTGGACC 21144
Qy 657 -----ValGlyPheGlyLysSerAlaProGlySerSerTyrProAlav 671
Db 21145 CCCCTCCATGCGCTTGCAGGAGCTGTTCGATTCCTTCCCTGCGCGGCT---CCGCGG 21201
Qy 671 alPheValValGlyThrIleGlyGlyValThrGlyAlaTyrArgSerAspAspCysGlyT 691
Db 21202 ATGCATCCAGTGGCAGCGCAATCTCGGCGCAGGAGGAGAAAGGCGGTGTGGG 21261
Qy 691 hrThrTrpValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnTrpGlyAlaIleT 711
Db 21262 TGCTC-----TCCACGG 21273
Qy 711 hrGlyAspHisAlaAsnLeuArgArgValTyrIleGlyThrAsnGlyArgGlyIleValT 731
Db 21274 CTGAGAGGGGCGACCTCCCTAGG-----GGAGAAGAGGCGACGTTGGGGGTTTC 21324
Qy 731 yrGlyAspIleGlyGlyAla 737
Db 21325 CGGGGGCGGGGGGAGCA 21344

```

RESULT 13

US-09-949-016-12639

; Sequence 12639, Application US/09949016

; Patent No. 6812339

```

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12639
; LENGTH: 37802
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(37802)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-12639

```

Alignment Scores:

```

Pred. No.: 0.00623 Length: 37802
Score: 177.00 Matches: 189
Percent Similarity: 28.8% Conservative: 56
Best Local Similarity: 22.2% Mismatches: 292
Query Match: 4.4% Indels: 314
DB: 3 Gaps: 48

```

US-09-917-376-3 (1-740) x US-09-949-016-12639 (1-37802)

```

Qy 14 GlyGlyGlyGlyPheValAspGly-----IleValPheAsnGlyGly 27
Db 19367 GGGGGCTCTGGCAGGGCGGACGCGTGGCTCCCTTCTTCACGGTTTATTCACAGGGGAC 19426
Qy 28 AlaProGlyIleLeuTyrValArgThr-----AspIleGlyGlyMet 41
Db 19427 AGGCTGGGAGTTGTTATTTGGCGCGCTGTTGGTGGCTGAGGGTGCAGGGACTT-GGGGG- 19482
Qy 42 TyrArgTrpAspAlaAlaAsn----- 48
Db 19483 TGGCGTGGGAGCGCGGAGGTATTAACGTATAATCATAGTAACACATCAGAAATG 19542
Qy 49 -----GlyArgTrpIleProLeu-LeuAspTrpValGlyTrpAsnAsnTrpGlyTyrAs 66
Db 19543 GACCCGAGCGCTGTCGCGCTAGCTCCAGCTCTCCCTGGCCCGCCGAGGAGA 19602
Qy 66 nGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaValG 86
Db 19603 GGGGTCCGATCCCTCCGCGGTCTCTCTCTC----- 19633
Qy 86 yMetTyrThrAsnSerTrpAspProAsnAspGlyAla-----IleLeuArgSerSe 103
Db 19634 -----CTGGTACCTGGCTTGGGTGGGGAACGAGCCTACTTCTGTACCGCTTT 19686
Qy 103 rAspGlnGlyAlaThr-----TrpGlnI 111
Db 19687 TGGCGAGCGGGGAGCCAGTGAATAGCCGTGGAGCCCGAGGCTGCTGGCTTTG 19746
Qy 111 eThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArgGlyMetGly----- 128
Db 19747 CGCACCGAGTCTTGG-----GGACCTGGTGTCCCGGAAACAACTTGGGACCTGGTA 19800
Qy 129 -----GluArgLeuAlaValAspProAsnAsnAsnIleLeuTyr----- 142
Db 19801 TCCCGGAGAGGCTTGGGACCTGGTGTCTCCGAGAGAGGCTTGGGTACCTGTTCTCT 19860
Qy 143 -----PheGlyAlaProSerGlyLysGlyLeu----- 151

```


19861 GGAAGAGCTTGGACACCTGGTCTCTGGGAGGCGCTTTGGACCTGGTCTCTGGGAGA 19920
152 ----TTPArgSerThrAspSerGlyAlaThrTTPSerGlnMetThrAsnProAspVa 170
19921 GCGTTGGAGATCTGTCTCTGGGAGAGGCTTGGGA----- 19957
170 lGlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGlnSerAspIleGlnGlyVa 190
19958 ----- 19974
190 lValTTPValAlaPheAspLysSerSerSerLeuGlyGlnAlaSerLysThrIlePh 210
19975 GGCITGG-----GACCTGGTGACCTTGGAGAGCTTGGACCTGGTGT 20019
210 eValGlyValAlaAspProAsnProValPheTTPSerArg----- 224
20020 TCTGGGAGAGCT-----TGGGACCTGGTGTCTCTGGGAGAGGCTTGGGACCTGGTGT 20073
225 -----AspGlyGlyAlaThrTTP-----GlnAlaValProG 235
20074 CTCTGGAAGAGCTTGGACACCTGGTGACCGCGGAGGCGCTTGGGATCTGGTGTCCCGG 20133
235 lValAlaProThrGlyPheIleProHisLysGlyValPheAspProValAsnHisValLeuT 255
20134 GAGACCTTGGGA-----CCTGTGTCTCTG----- 20159
255 yRIleAlaThrSerAsnThrGlyGlyProTyrAspGlySerSerGlyAspValTTPlysp 275
20160 -----GGAGAGCTTGGGACCT-----GGTGACCTTGGAGAGCTTGGGAC 20202
275 heSerValThrSerGlyThrTTPThrArgIleSerProValProSerThrAspThrAlaA 295
20203 CTGGTGTCTCTGAGAGAGCTTGGGA-----TCTGGTGTCTCCAGGAGAGCTTGGGAC 20256
295 snAspTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetV 315
20257 CT-----GGTGTCTCTGGAAGAGCTTGGAC----- 20282
315 alAlaThrGlnIleSerTTPPro-----AspThrIleIlePheArgs 330
20283 -----ACCTGTGTCTCTGGGAGAGCTTGGGACCTGGTG----- 20318
330 erThrAspGlyGlyAlaThrTTPThrArgIleTTPAsp-TTPThrSerTyrProAsnArg 349
20319 -----TCTCTGGGAGAGCTTGGGACCTGTGT-----CCTGGGAGA 20355
350 SerLeu-ArgTyrValLeuAspIleSerAlaGluProTTPLeuThrPheGlyValGlnPr 369
20356 GCGTTGGAGA-----TCTGTGTAGCCGGGAGAGCTTGGGA----- 20392
369 oAsnProProValProSerProLysLeuGlyTTPMetAspGluAlaMetAlaIleAspPr 389
20393 -----CCTGGTGTCTCCGGGAGAGCTTGGGACCTGGTGTCTCCGGGAGAGCTTGGACACC 20448
389 o-PheAsnSerAspArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspL 409
20449 TGGTGTCTCCAGGAGAGCTT-----GGGACCTGGTGACCTTGGAGAGGCTGGGAC 20502
409 euThrLysTTPAspSerGlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluG 429
20503 TGGTGACCCGGGAGAGCTTGGG----- 20525
429 luThrAlaValAsnAspLeuIleSer-----ProProSerGlyAlaProLeuIleSerA 447
20526 -----GACCTGTGTCTCTGGGAGAGCTTGGGACCT----- 20558
447 laLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThrAlaValProSerThrIleP 467
20559 -----GGTGACCTTGGAGAGCTTGGGACCTGGTGTCTCGGAGTGCCTTGGGACCT- 20612
467 heThrSerProValPheThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerI 487
20613 --AGTGACCCGGGAGAGCTTGGGACCTGTGTCTCCGGGAGAGGCTTGGGACCTGTGT 20670

QY 487 leIleValArgAlaGlySerPheAspProSerSerGlnProAsnAspArgHisValAlap 507
Db 20671 TCCTG-----GGAGAGCTTGGGACCTGGTGTCTCTG- 20702
QY 507 heSerThrAspGlyGlyLysAsnTTPPheGlnGlySerGluProGlyGlyValThrThrG 527
Db 20703 -----GGAGAGGCTTGGGACCTGTGTCTCGGAGAGAGCTTGGGACCTGGTGACC 20757
QY 527 lyGlyThrVal-----AlaAlaSerAlaAspGlySerArgPheValTTPAlaProGlyAspP 546
Db 20758 CGGAGAGGCTTGGACACCTGTGTCTCCGGGAGAG-----GCTTGGGACCTGGTGACC 20811
QY 546 roGlyGln-----ProValValTyrAlaValGlyPheGlyAsnSerTTPAlaAlas 563
Db 20812 CGGAGAGGCTTGGGACCTGTGTCTCGGAGAGAGCTTGGGG-----ACCTGTGTCTCG 20868
QY 563 erGln-----GlyValProAlaAsnAlaGlnIleArgSerAspArgVala 578
Db 20869 CGAGAGAGCTTGGGACCTGTGTGACCCGGGAGAGCTTGG----- 20909
QY 578 snProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGly--- 596
Db 20910 -----ACACCTGGTGTCTC 20922
QY 597 --GlyValThrPheGlnProValAlaAlaGlyLeuPro---SerSerGlyAlaValGlyV 615
Db 20923 CGGAGAGGCTTGGGACCTGTGTCTCCGGGAGAGCTTGGGACCTGGTGTGAG 20982
QY 615 alMetPheHisAlaValProGlyLysGlyGlyAspLeuTTPLeuAlaAlaSerSerGlyL 635
Db 20983 AGGCTTGGGA-----CCTGTGTATCTTGGAGAGGCTTGGGACCTGTGTCTCGGA- 21035
QY 635 euTyrHisSerThrAsnGlyGlySerSerTTPSerAlaIleThrGlyValSerSerAlav 655
Db 21036 -----GAGTTACGGGGCTGTGTGGGGAGAGACCTTGTGAGCAAAAGTCC 21084
QY 655 alaSn----- 656
Db 21085 TGAATCCCTCGAAGAGAGCGCATCGGAGCTCCCTCGAGGCGTTCATTTGTGGACC 21144
QY 657 -----ValGlyPheGlyLysSerAlaProGlySerSerTyrProAlav 671
Db 21145 CCCCTCCCATCGCTTTCAGGAGAGCTGTTCGATTCCTCCCTGCGCGGCT---CCGCGG 21201
QY 671 alPheValValGlyThrIleGlyGlyValThrGlyAlaTyrArgSerAspCysGlyT 691
Db 21202 ATGCATCCAGTGGCAGCCCAATTCTGGGCCAGGGGAGAGGAAAGCGGGTGTGGGG 21261
QY 691 hrThrTTPValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnTTPGlyGlnAlaIleT 711
Db 21262 TGGTC-----TCCACGG 21273
QY 711 hrGlyAspHisAlaAsnLeuArgArgValTyrIleGlyThrAsnGlyArgGlyIleValT 731
Db 21274 CTGGAGAGGGCGAGCTTCCCTAGG-----GGAGAAGAGGACCGTGGGGGTTC 21324
QY 731 yrGlyAspIleGlyGlyAla 737
Db 21325 CGGGGGCGGGCGGAGCA 21344

RESULT 14

US-09-902-540-2590
; Sequence 2590, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540

<hr/>					
; CURRENT FILING DATE: 2001-07-10					
; PRIOR APPLICATION NUMBER: 60/217,883					
; PRIORITY FILING DATE: 2000-07-10					
; NUMBER OF SEQ ID NOS: 16825					
; SEQ ID NO 2590					
; LENGTH: 4287					
; TYPE: DNA					
; ORGANISM: Myxococcus xanthus					
US-09-917-376-3 (1-740) x US-09-902-540-2590 (1-4287)					
Alignment Scores:					
Pred. No.:	3,07e-05	Length:	4287		
Score:	175.50	Matches:	165		
Percent Similarity:	35.3%	Conservative:	90		
Best Local Similarity:	22.8%	Mismatches:	240		
Query Match:	4.3%	Indels:	229		
DB:	3	Gaps:	47		
US-09-917-376-3 (1-740) x US-09-902-540-2590 (1-4287)					
QY	24	PheAsnGluGlyAlaProGlyLeuTyrValArgThrAspIleGlyMetTyrArg	43		
Db	2299	TACAACCTCAGTTCCGGGCATGCTCCCTCGGCACCATCGACTGGGCACGCCCTACGCC	2358		
QY	44	TrpAspAlaAAsnGlyArgTrpIleProLeuLeuAsp-----TrpValGly	59		
Db	2359	GGCAAGACGGTCAGTGCCGTTCGGGGTCGGGTACGACGAGAGCGCGGCTACACGGGC	2418		
QY	60	Trp-----AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsn	77		
Db	2419	TGGTGTGGACGACCTCGAGTTCAACGGCATC-----	2451		
QY	78	ThrAsnLysValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGly	97		
Db	2452	ACGAACACTCCGTTCGCCACCATCGCG-----CCCGAGGATGCG	2490		
QY	98	AlaIleLeuArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLys	117		
Db	2491	GTTTGT-----GTGAACCCCTGGGCCCATCGCC	2517		
QY	118	LeuGlyGlyAsnMetProGlyArgGlyMet-----GlyGluArgLeu-----	131		
Db	2518	AACGGCGGT-----CCGGACCGGTCCATCGCGCGGGTGAACCTGTGAGCCTCTACGGT	2571		
QY	132	---AlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGly	150		
Db	2572	TCGGCTGCCATCCGGAAGCCCGCGGTGACCTTC-----	2607		
QY	151	LeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPheProaspVal	170		
Db	2608	-----ACCTGGGCACAG-----ACGTCTGGCCCTCGCGTC	2637		
QY	171	----GlyThrTyrlleAlaAsnProThrAspThrGlyTyrGlnSerAspIle	187		
Db	2638	ACCCTGGCGGGTCCNACACCTGAACCCGTCTTCACGGCGCGCGAGGTACCCGAGTCC	2697		
QY	188	GlnGlyValValTrpValAlaPheAspLysSerSerSerSerLeuGlyGlnAlaSerLys	207		
Db	2698	ACCGCACTGGTCTTCACGGTG-----ACCGTCTCGGCGCGGTCAAGACGTCCAGGNAC	2751		
QY	208	ThrIlePheValGlyValAlaAspProAsnAsnProValPheTrpSerArgAspGlyGly	227		
Db	2752	TCGGTGAACCGTGGCGGTCTCCCAACAAATCCG-----	2787		
QY	228	AlaThrTrpGlnAlaValProGlyAlaProThrGlyPheIleProHisLysGlyValPhe	247		
Db	2788	-----CCCACCGTGAACCGGGCCTTCGACGGCATCGTCGAGGAGCGCGCTGAGTAC	2838		
QY	248	AspProValAsnHisValLeuTyrlleAlaThrSerAsnThrGlyGlyProTyrrAspGly	267		
Db	2839	-----ACGCTGAGCGGCTCCGCCACGATCGGATGCG	2871		
QY	268	SerSerGlyAspValTrpLysPheSerValThrSerGlyThrTrpThrArgIleSerPro	287		
<hr/>					
Db	2872	AATGCC-----CTCACGTACCTGTGGAGCCAGGTCTTCGGT	2907		
QY	288	ValProSerThrAspThrAlaAsnAspTyrPheGlyTyrSerGlyLeuThrIleAspArg	307		
Db	2908	ACGCCGGTG-----CCGGTGAAGGACTAC-----	2931		
QY	308	GlnHisProAsnThrIleMetValAlaThrGlnIleSerTrpTrpProAspThrIleIle	327		
Db	2932	ACCACGCGAGCGGACGTTTCATCGCGCCGAGGTACAG-----CTGGATGAGTGGCTGTC	2988		
QY	328	PheArgSerThr-----AspGlyGlyAlaThrTrp-ThrArgIleTrpAspTrpThrSe	345		
Db	2989	TTCCGCTCAGCGTCAGCGACGCGCATCGCACCGGTGAACG-----ACACG	3033		
QY	345	rTyPrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGluProTrpLeuThrPh	365		
Db	3034	GTACACGTGACGTTCACCAACGCCA-----ACCGCGGCCCA-----	3070		
QY	365	eGlyValGlnProAsnProProValProSerProLysLeuGlyTrpMetAspGluAlaMe	385		
Db	3071	----TCGTCTCCGACACGCTCCGTGGCTTCGCGCGGGCACT-GTCACCGTGACGGCTC	3125		
QY	385	tAlaIleAspPropHeAsnSerAspArgMetLeuTyrGly-----ThrGlyAlaTh	402		
Db	3126	CGCGGTGCAACCG-----GATGGTGACGCGCTCAGCTACAGCTGGGAGCAGACGGGTGTT	3182		
QY	402	rLeuTyrrAlaThrAsn-----	407		
Db	3183	GACGGTGGCCATCAATGGCGCGACAGCTCGGCCCATCTCTTCGCCACCCCGGTCCCGG	3242		
QY	408	-----As	408		
Db	3243	TTCTACGAGTTTCACCGTGACGCGGACGCGCTCGGCCCTCTCCCTCCAAGCGGTGCC	3302		
QY	408	pLeuThrLysTrpAspSerGlyGlyGlnIleHisIleAlaProMetVallys---GlyLe	427		
Db	3303	GGTGACCATCATCGATGGTTCTCTGCGCGGCAACTCTGCGCCGACCTGTGAACGCGGCAT	3362		
QY	427	uGluGluThrAla---ValAsnAspLeuIleSerProProSerGlyAlaProLeuIleSe	446		
Db	3363	CGATGCCACCGGGAACCGCTGGTGACGTCTGTGACGTGACGCTCAGCGGTCCGCGAGC	3416		
QY	446	rAlaLeuGlyAsp-----LeuGlyGly-----Ph	454		
Db	3417	CGCCGAGGTGACACGCTGACGTACCATCGGAGCAGATCGGTGGCAGCGATGTGACCT	3476		
QY	454	eThrHisAlaAspValThrAlaValProSerThrIlePheThrSerProValPheThrTh	474		
Db	3477	GACGGCGCGCGAC-----ACCTGACGCGCTCG-----TTCAACGCGCGGACGACGCGCAG	35		

```

Db 3795 GCCGAGCTTACCGCCCGGCTCTCTCC-----GATCGCTGAC 3833
Qy 581 rPheTyrAlaLeu---SerAsnGly-----ThrPheTyrArgSerThrAspGlyGlyVa 598
Db 3834 CTTCTCTCTGATGTCAGCGAGCGGTACGCGCACTGTCTGACGTGACGTCCATCAGCGT 3893
Qy 598 lThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHi 618
Db 3894 GACCAGGAGAGACTCGCG-----CCGGTGGCTGCCGTCTCGCGGGTGCTC----- 3939
Qy 618 sAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSe 638
Db 3940 -----TCTGCAACACGACGCTCGGCCACGCTCGACGGCTCGGGTTCAG 3983
Qy 638 rThrAsnGlyGlySer-----SerTrpSerAlaAlaThrGlyValSerSerAl 654
Db 3984 CGACCGAAACGGCGATGTGCTGACGTACCGCTGGACGCGAGTCTCGCGCCCGAAGCGCAC 4043
Qy 654 aValAsn 656
Db 4044 CATCTCC 4050

RESULT 15
US-09-917-376-3 (1-740) x US-09-902-540-1119 (1-16584)
; Sequence 1119, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1119
; LENGTH: 16584
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1119

Alignment Scores:
Pred. No.: 0.000242 Length: 16584
Score: 175.50 Matches: 165
Percent Similarity: 35.3% Conservative: 90
Best Local Similarity: 22.8% Mismatches: 240
Query Match: 4.3% Indels: 229
DB: 3 Gaps: 47

US-09-917-376-3 (1-740) x US-09-902-540-1119 (1-16584)
Qy 24 PheAsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyMetTyrArg 43
Db 6448 TACAACCTTCAGTTCGCGGATCTCCCTCGACCATCGACCTGGCGCGCCATCGCC 6507
Qy 44 TrpAspAlaAlaAsnGlyArgTrpIleProLeuLeuAsp-----TrpValGly 59
Db 6508 GCGAAGACGGTGCAGGTCCGCTTCGGGTCCGTTACGGACGAGAGCGCGGCTACACGCGC 6567
Qy 60 Trp-----AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsn 77
Db 6568 TGGTGTGGACGACCTCGAGTTCAACGGCATC----- 6600
Qy 78 ThrAsnLysValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGly 97
Db 6601 ACGAAGACTCGTTCGCGCACCATCGCG-----CCGAGGATGGC 6639
Qy 98 AlaIleLeuArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLys 117
Db 6640 GTTTGT-----GTGAACCCCTTGCCCCATCGCC 6666

```

```

Qy 118 LeuGlyGlyAsnMetProGlyArgGlyMet-----GlyGluArgLeu----- 131
Db 6667 AACGCGGT-----CCGACCGGTCCATTCGCGCGGGTGAATGCTGTGAGCCTTACGGT 6720
Qy 132 ---AlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGly 150
Db 6721 TCGGTGCGGATCCGGAAGCCGCGGTGACCTTC----- 6756
Qy 151 LeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPheProAspVal 170
Db 6757 -----ACCTGGACAG-----ACGTCTGGCCCTGCGGTC 6786
Qy 171 -----GlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGlnSerAspIle 187
Db 6787 ACCCTGGGGTGCACACACCTGAACCGTTCGCTTACGCGCGCGCAGGTACCGAGTCC 6846
Qy 188 GlnGlyValValTrpValAlaPheAspLysSerSerSerLeuGlyGlnAlaSerLys 207
Db 6847 ACCGCACTGGTGTTCACGCTG-----ACCGTCTCCGACGGGTCAAGACGTCCACGGAC 6900
Qy 208 ThrIlePheValGlyValAlaAspProAsnAsnProValPheTrpSerArgAspGlyGly 227
Db 6901 TCGGTGACGTGACCGTGGCGCTCCCAACAATCCG----- 6936
Qy 228 AlaThrTrpGlnAlaValProGlyAlaProThrGlyPheIleProHisLysGlyValPhe 247
Db 6937 -----CCACCGTGAACGCGGCGCTCGACGGCATCGTCGAGGAGCGCGTGTAGTAC 6987
Qy 248 AspProValAsnHisValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyrAspGly 267
Db 6988 -----ACGTGACGCGGCTCCGCCAGCATGCGGATGCG 7020
Qy 268 SerSerGlyAspValTrpLysPheSerValThrSerGlyThrTrpThrArgIleSerPro 287
Db 7021 AATGCC-----CTCAGTACTGTGGACCCAGGTCTCCGGT 7056
Qy 288 ValProSerThrAspThrAlaAsnAspTyrPheGlyTyrSerGlyLeuThrIleAspArg 307
Db 7057 ACGCGGTG-----GCGGTGAGGACTAC----- 7080
Qy 308 GlnHisProAsnThrIleMetValAlaThrGlnIleSerTrpTrpProAspThrIleIle 327
Db 7081 ACCACGCGGCGGCGACGTTTCATCGCGCGAGGTCAAG-----CTGGATCAGTCTCGTGGTC 7137
Qy 328 PheArgSerThr-----AspGlyGlyAlaThrTrp---ThrArgIleTrpAspTrpHse 345
Db 7138 TTCGTCTGACGCGTCAAGCGCATCGCCACGTTGAACG-----ACACG 7182
Qy 345 rTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGluProTrpLeuThrPh 365
Db 7183 GTCACCGTACGCGTCAACACCGCA-----ACCGCGCGCCCA----- 7219
Qy 365 eGlyValGlnProAsnProProValProSerProLysLeuGlyTrpMetAspGluAlaMe 385
Db 7220 -----TCGTCTCGACACGTCGTCGCTTCGCGCGGGCACT-GTCACCGTACGCGCTC 7274
Qy 385 tAlaIleAspProPheAsnSerAspArgMetLeuTyrGly-----ThrGlyAlaTh 402
Db 7275 CGCGTTCGACCCG---GATGGTACGCGCTCAGCTACAGCTGGGAGCAGACGCGGTGTC 7331
Qy 402 rLeuTyrAlaThrAsn----- 407
Db 7332 GACGTGCGCCATCAATGCGCGCACACGTCCGCCATCTCTTTCGCCACCCCGGTCCCGGG 7391
Qy 408 -----As 408
Db 7392 TTCGTACAGTTCACCGTACGCGCGACGCGGCTCGGCTCTGCTCTCAAGCGGTGTC 7451
Qy 408 pLeuThrLysTrpAspSerGlyGlyGlnIleHisIleAlaProMetValLys---GlyLe 427
Db 7452 GGTGACCATCATCGATGTTCTCTCGCGCGCAACTCTGCGCGCGACGTCGTGAACGCGGCGCAT 7511

```

```
QY 427 uGluGluThrAla---ValAsnAspLeuIleSerProProSerGlyAlaProLeuIleSe 446
Db 7512 CGATGCCACGCGGAACGCTGGTGACGTGACGCTCAGCGGCTCCGCGAGC-----GA 7565
QY 446 rAlaLeuGlyAsp-----LeuGlyGly-----Ph 454
Db 7566 CGCGAGGGTGACACGCTGACGTACCACTGGGACGACATCGTGGCAGCGATGTGACCT 7625
QY 454 eThrHisAlaAspValThrAlaValProSerThrIlePheThrSerProValPheThrTh 474
Db 7626 GACGGGCGCGAC---ACCCTGACGCGCTCG-----TTACCGCGCGGACGACGCGCAG 7676
QY 474 rGlyThrSerValAspTyrAlaGluLeuAsnProSer----- 486
Db 7677 CGGTACCACGCTCGGCTTCATC---CTGACGGTACGGATGGCACCTCGACCCACGACGA 7733
QY 487 -----IleIleValArgAlaGlySerPheAspProSerSerGlnProAsnAspAr 503
Db 7734 TGCGGTGCGGTCACTGTCGGCGG-----GACCCGGTGCCCAACCCGAGC----- 7779
QY 503 gHisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGl 523
Db 7780 -AACACTGCTCCGGAAGCGAGCGGCGAGTCGCGCATCGTG---GCGAGGGCGGCCAC 7835
QY 523 yValThrThrGlyGlyThrValAlaAlaSerAlaAspGlySerArg-----PheValTr 541
Db 7836 TGTGACGCTCAACGCGCAG---GCGACCGACGCTGACGGTGACACGCTCGTCACTGCTG 7892
QY 541 pAlaProGlyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrpAl 561
Db 7893 GAGCGAG---ATCGGCGGACGCGCCGCG---ACGCTGTCTGACGCGCTCTCGCTGAC 7943
QY 561 aAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysTh 581
Db 7944 GCCGACGTTACCGCCCGGCTCTCTC-----GATCCGCTGAC 7982
QY 581 rPheTyrAlaLeu---SerAsnGly-----ThrPheTyrArgSerThrAspGlyGlyVa 598
Db 7983 CTTCTCTCTGATGTCAGCGACGCTACGGCGACTGTCGTGACGTGACGTCCATCAGCGT 8042
QY 598 lThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHi 618
Db 8043 GACCGAGGAGAACGTCGCG-----CCGGTGGCGTCCGCTCGCGGTGCTC----- 8088
QY 618 sAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSe 638
Db 8089 -----TCTGGCACACGACGTCGCGCCACGCTCGACGCGCTCGGCTTCCAG 8132
QY 638 rThrAsnGlyGlySer-----SerTrpSerAlaIleThrGlyValSerSerAl 654
Db 8133 CGACGGAACGGCGATGTCGTGACGTACCGTGGACCGAGGTCTCCGGCCCGACGCGAC 8192
QY 654 aValAsn 656
Db 8193 CATCTCC 8199
```

Search completed: March 2, 2006, 13:50:24
Job time : 5618.94 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 2, 2006, 12:19:26 ; Search time 1388.78 Seconds
(without alignments)
4406.259 Million cell updates/sec

Title: US-09-917-376-3

Perfect score: 4036

Sequence: 1 ATTQPTWSNAIGGGGFD.....YGTNGRGIVGIGGAPSG 740

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abss/ABSSWEB.spool/US09917376/runat_02032006_091500_8223/app_query.fasta_1
-DB=Published Applications_NA_Main -QMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss06p
-USER=US09917376 @CGN_1_1_3095 @runat_02032006_091500_8223 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_NA_Main:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4036	100.0	2869	3	US-09-917-376-2
2	4036	100.0	2869	5	US-10-155-400-2
3	2429.5	60.2	2646	6	US-10-156-761-1845
4	2429.5	60.2	9025608	6	US-10-156-761-1
5	1579	39.1	2517	5	US-10-026-994-4
6	1579	39.1	2710	5	US-10-026-994-1
7	1405	34.8	2217	6	US-10-156-761-2561
					Sequence 2, Appli
					Sequence 2, Appli
					Sequence 1845, Ap
					Sequence 1, Appli
					Sequence 4, Appli
					Sequence 1, Appli
					Sequence 2561, Ap

Alignment Scores:
Pred. No.: 0
Score: 4036.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Length: 2869
Matches: 740
Conservative: 0
Mismatches: 0

c	8	1405	34.8	9025608	6	US-10-156-761-1	Sequence 1, Appli
	9	1139	28.2	3668	3	US-09-927-827-21	Sequence 21, Appli
	10	1117	27.7	2367	7	US-10-395-241-13	Sequence 13, Appli
	11	1117	27.7	2646	7	US-10-395-241-11	Sequence 11, Appli
	12	1114	27.6	2481	7	US-10-395-241-17	Sequence 17, Appli
	13	878.5	21.8	5698	7	US-10-420-191-1	Sequence 1, Appli
	14	392.5	9.7	1103	8	US-10-653-047-7511	Sequence 7511, Ap
	15	238	5.9	7407	6	US-10-246-330-3	Sequence 3, Appli
	16	238	5.9	7407	7	US-10-282-122A-30151	Sequence 30151, A
	17	199.5	4.9	7305	7	US-10-282-122A-11269	Sequence 11269, A
	18	195	4.8	5295	7	US-10-282-122A-11871	Sequence 11871, A
	19	192	4.8	9903	7	US-10-282-122A-28185	Sequence 28185, A
	20	188.5	4.7	4399	6	US-10-369-493-37642	Sequence 37642, A
	21	184	4.6	3324	7	US-10-282-122A-26288	Sequence 26288, A
	22	182	4.5	69350	7	US-10-041-018-379	Sequence 379, App
	23	180	4.5	2355	6	US-10-369-493-37792	Sequence 37792, A
	24	179.5	4.4	3222	7	US-10-282-122A-14432	Sequence 14432, A
	25	176	4.4	3267	5	US-10-158-160A-14	Sequence 14, Appli
	26	174	4.3	2855	10	US-11-097-143-40748	Sequence 40748, A
	27	174	4.3	5221	10	US-11-097-143-26683	Sequence 26683, A
	28	174	4.3	5302	10	US-11-097-143-40747	Sequence 40747, A
	29	173	4.3	2708	10	US-11-097-143-26684	Sequence 26684, A
	30	170.5	4.2	2176	10	US-11-097-143-14408	Sequence 14408, A
	31	170.5	4.2	4252	10	US-11-097-143-14407	Sequence 14407, A
	32	170.5	4.2	7191	7	US-10-282-122A-35048	Sequence 35048, A
	33	170.5	4.2	25165	5	US-10-114-170-39	Sequence 39, Appli
	34	170	4.2	4323	7	US-10-282-122A-8383	Sequence 8383, Ap
	35	169.5	4.2	3360	6	US-10-245-802-11	Sequence 11, Appli
	36	169	4.2	69350	7	US-10-041-018-379	Sequence 379, App
	37	168	4.2	5451	6	US-10-101-510-597	Sequence 597, App
	38	167.5	4.2	2110	5	US-10-027-632-97984	Sequence 97984, A
	39	167.5	4.2	2110	6	US-10-027-632-97984	Sequence 97984, A
	40	167	4.1	1839	6	US-10-369-493-34868	Sequence 34868, A
	41	166	4.1	3879	7	US-10-282-122A-8717	Sequence 8717, Ap
	42	165.5	4.1	4383	6	US-10-245-802-21	Sequence 21, Appli
	43	165	4.1	6615	7	US-10-282-122A-28180	Sequence 28180, A
	44	165	4.1	67167	9	US-10-937-730A-3	Sequence 3, Appli
	45	164.5	4.1	2540	9	US-10-680-860A-161	Sequence 161, App

ALIGNMENTS

RESULT 1

US-09-917-376-2
; Sequence 2, Application US/09917376
; Publication No. US20040038334A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOYTICUS
; FILE REFERENCE: 40197.4US01
; CURRENT APPLICATION NUMBER: US/09/917,376
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2869
; TYPE: DNA
; ORGANISM: Acidothermus cellulolyticus
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (2869)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-917-376-2

Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0
US-09-917-376-3 (1-740) x US-09-917-376-2 (1-2869)

QY 1 AlaThrThrGlnProTyrThrTrpSerAsnValAlaIleGlyGlyGlyPheValAsp 20
DB 139 GCGACGACTCAGCCGTACACCTGGAGCAACGTGGCGATCGGGGCGCGCTTTGCGAC 138
QY 21 GlyIleValPheAsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGly 40
DB 199 GCGATCGCTCTCAATGAAGGTGCACCGGAATCTGTACGTGCGGACGACATCGGGGG 258
QY 41 MetTyrArgTrpAspAlaIleAlaGlyArgTrpIleProLeuLeuLeuAspTrpValGlyTrp 60
DB 259 ATGATTCGATGGATGCGCGCAACCGCGGTGGATCCCTCTTCGTGGATTGGTGGATGG 318
QY 61 AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLys 80
DB 319 AACAAITGGGGGTACACCGCGTGTGTCAGCATTTGCGGACACCGATCAATCAACAAG 378
QY 81 ValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeu 100
DB 379 GTATGGCGCCGCTCGGAATGTACACCAACAGCTGGGACCCCAACCGAGCGATTCTC 438
QY 101 ArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGly 120
DB 439 CGCTCGTCTGATCAGGGCGCAACGTGGCAATAAATCGCCCTGCGGTTCAAGCTTGGCGG 498
QY 121 AsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIle 140
DB 499 AACATCCCGGGGGTGGAAATGGCGAGCGGCTTGGCGTGGATCCAAACAATGACACATT 558
QY 141 LeuTyrPheGlyAlaProSerGlyLysGlyLeuTyrArgSerThrAspSerGlyAlaThr 160
DB 559 CTGTATTTTGGCGCCCGAGCGGCAAGGGCTCTGGAGAGCACACAGATTCCGCGCGGACC 618
QY 161 TrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThr 180
DB 619 TGTCTCCAGATGACGAATTTTCGGACGTAGGCACGTACATTGCAAAATCCCACTGACACG 678
QY 181 ThrGlyTyrGlnSerAspIleGlnIleValValTrpValAlaPheAspLysSerSer 200
DB 679 ACCGGCTATCAGAGCGATATTCAAGCGCTCGTCTGGGTGCGCTTTTCAGCAAGTCTTCGTCA 738
QY 201 SerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspProAsnAsnProVal 220
DB 739 TCGTCTGGCGAAGCGAGTAAAGACCAATTTTGTGGCGTGGCGATCCCAATAATCCGGTC 798
QY 221 PheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPhe 240
DB 799 TTCTGGAGCAGACAGCGCGCGGACGTGGCAGCGGTGCGGGTGGCGGCGGCTTC 858
QY 241 IleProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsn 260
DB 859 ATCCCGCACAAAGGGGCTCTTTGACCCGGTCAACACAGTGTCTATATTGCCACACGAAT 918
QY 261 ThrGlyGlyProTyrAspGlySerSerGlyAspValTrpLysPheSerValThrSerGly 280
DB 919 ACGGGTGGTTCGATGACGGAGCTCCGGGACGTCTGGAAATTCCTGGGTGACCTCCGGG 978
QY 281 ThrTrpThrArgIleSerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyr 300
DB 979 ACATGGACCGGAATAGCCCGGTACCTTCAGCGGACACCGGCCAACGACTACTTTGGTTAC 1038
QY 301 SerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSer 320
DB 1039 AGCGGCTCACTATGACCGCGCAGCACCCGCAACAGATTAATGGTGGCAACCCAGATATCG 1098
QY 321 TrpTrpProAspThrIleIlePheArgSerThrAspGlyGlyAlaThrTrpThrArgIle 340
DB 1099 TGGTGGCGCGGACCACTAATCTTCGGAGCACCGCGGGTGGAGCTGGAGCGCGGATC 1158

QY 341 TrpAspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGlu 360
DB 1159 TGGATTGGACGAGATTATCCCAATTCGAAGCTTGGGATATGCTGTGACATTTCCGCGGAG 1218
QY 361 ProTrpLeuThrPheGlyValGlnProAsnProProValProSerProLysLeuGlyTrp 380
DB 1219 CTTTGGTGTACCTTCGGCGTACAGCGAATCTCCGTACCCAGTCCAGTCCGAGCTCGGCTGG 1278
QY 381 MetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThrGly 400
DB 1279 ATGGATGAACGATGGCAATCGATCCGTTCAACTCTGATCGATGCTCTACGGAACAGC 1338
QY 401 AlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlyGlnIleHisIle 420
DB 1339 GCGACGTTGTACGCAACAATGATCTCAGAAAGTGGGACTCCGCGCGGCGAGATTCATATC 1398
QY 421 AlaProMetValLysGlyLeuGluGluThrAlaValAsnAspLeuIleSerProSer 440
DB 1399 GCGCGGATGGTCAAGGATTGGAGAGCGCGGTAAACGATCTCATCAGCCCGCGCT 1458
QY 441 GlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThr 460
DB 1459 GCGCGCCCGCTCATCAGCGCTCTCGGAGACTTCGCGGCTTCACCACGCGGAGCTTACT 1518
QY 461 AlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSerValAspTyr 480
DB 1519 GCGGTGCTCATCGACGATCTTCACTCACCGTGTTCACGACCGGCGGACCGGCTGACTAT 1578
QY 481 AlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspProSerSerGlnPro 500
DB 1579 GCGGAATTGAATCCGTCGATCATCGTTGCGGTGGAAGTTTCGATCCATCGAGCCAAACG 1638
QY 501 AsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGlySerGlu 520
DB 1639 AACGACAGGACAGCTCGGCTTCTCGACAGAGCGGCGCAAGAACTGGTTCGAAGCAGCGAA 1698
QY 521 ProGlyGlyValThrThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheVal 540
DB 1699 CCTGGCGGGGTGACGACGCGGCGCACGTCGCGCATCGGCGGACGGCTCTCGTTTCTGTC 1758
QY 541 TrpAlaProGlyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrp 560
DB 1759 TGGGCTCCCGCGCATCCCGTTCAGCTGTGTGTGTACGACGTGCGATTGGCACTCTCTGG 1818
QY 561 AlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLys 580
DB 1819 GCTGCTTCGAAAGGTGTTCGCGCAATGCCAGATCCGCTCAGACCGGGTGAATCCAAAG 1878
QY 581 ThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThrPhe 600
DB 1879 ACTTTCTATGCCCTATCCATGGNACCTTCTATCGAAGCAGCGGACGGGCGGTGACATT 1938
QY 601 GlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHisAlaVal 620
DB 1939 CAACCGGTCCGCGCGGTCTTCCAGCAGCGGTGCGCTCGGTGTGATGTTCCACCGCGTG 1998
QY 621 ProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsn 640
DB 1999 CTGGAAGAAAGAGCGGATCTGTGGCTCGCTGCAATCGAGCGGGCTTTTACCCTCAACCAAT 2058
QY 641 GlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnValGlyPheGly 660
DB 2059 GCGCGGACAGATTGGTCTGCAATCACCGGCTATCTCCGCGGTGAACGTGGGATTGGT 2118
QY 661 LysSerAlaProGlySerSerTrpProAlaValPheValValGlyThrIleGlyVal 680
DB 2119 AGTCTGCGCGCGGTGCTCATACCGCGCTCTTTGTCGTCGCGCAGATCGGAGCGGT 2178
QY 681 ThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAspAspGln 700
DB 2179 ACGGGGCGGTACCGCTCCGACGACTGTGGGACGACCTGGGTACTGTATCAATGATGACCA 2238
QY 701 HisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgVal 720

Db 2239 CACCAATACGGAATTTGGGACACCAATACCGGTGACCGCAATTTACGGCGGGTG 2298
Qy 721 TyrileGlyThrAsnGlyArgGlyIleValTyrGlyAspIleGlyGlyAlaProSerGly 740
Db 2299 TACATAGGACCAACGGCGGTGAATTCATACGGGAGCATTTGGTGGCGCGTCCGGA 2358

RESULT 2

US-10-155-400-2
; Sequence 2, Application US/10155400
; Publication NO. US20030108988A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOLYTICUS
; FILE REFERENCE: NREL 01-36A
; CURRENT APPLICATION NUMBER: US/10/155,400
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2869
; TYPE: DNA
; ORGANISM: Acidothermus cellulolyticus
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (2869)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-10-155-400-2

Alignment Scores:

Pred. No.: 0 Length: 2869
Score: 4036.00 Matches: 740
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-09-917-376-3 (1-740) x US-10-155-400-2 (1-2869)

Qy 1 AlaThrGlnProTyrThrTrpSerAsnValAlaIleGlyGlyPheValAsp 20
Db 139 GCGAGCTACCGCTACACCTGGAGCAACGTGGCGATCGGGCGCGGCTTTGTCCAC 198
Qy 21 GlyIleValPheAsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGly 40
Db 199 GGGATCGTCTTCAATGAAGGTGCACCGGAATTCGTACGTCCGACCGACATCGGGGG 258
Qy 41 MetTyrArgTrpAspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrp 60
Db 259 ATGTATCGATGGGATGCGCGCAACGGCGGTGGATCCCTCTTCGATGGGTGGGATGG 318
Qy 61 AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnIys 80
Db 319 AACAAATGGGGGTACAACGGCGTGTACAGATTCGGCAGACCCCGCATCAATACAAACAG 378
Qy 81 ValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeu 100
Db 379 GTATGGCGCGCGTCCGAATGTACCAACAGCTGGGACCCCAACAGCGGAGCATTCCTC 438
Qy 101 ArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGly 120
Db 439 CGCTCGTCTGATAGGCGCAACGTGGCAATAACGCCCTTCGCGTTCAAGCTTGGCGGC 498
Qy 121 AsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIle 140
Db 499 AACATGCGCGCGGTGAATGGCGAGCGGTTCGGTGGATCCAAACAAATGACAACAT 558
Qy 141 LeuTyrPheGlyAlaProSerGlyLysGlyLeuTyrArgSerThrAspSerGlyAlaThr 160

Db 559 CTGTATTTCCGGCGCCCGAGCGCAAAAGGCTCTGGAGAAGCACAGATTCCGGCGCGACC 618
Qy 161 TrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThr 180
Db 619 TGGTCCAGATGACGAATTTCCGGACGTAGGACAGTACATTGCAATCCCACTGACACG 678
Qy 181 ThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAspLysSerSerSer 200
Db 679 ACCGGCTATCAGAGCGATATTCAAGGCGTCTGTGGGTGCTTTCCGCAAGTCTTCGTCA 738
Qy 201 SerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspProAsnAsnProVal 220
Db 739 TCGCTCGGCAAGCGAGTAGTAGCAATTTTGTGGCGGTGGCGATCCCAATAATCCCGTC 798
Qy 221 PheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPhe 240
Db 799 TTCTGGAGCAGAGACGGCGCGCGAGCGTGGCAGCGGTGCGGTCGCCCGCTTC 858
Qy 241 IleProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsn 260
Db 859 ATCCCGCACAAAGGGCGTCTTTGACCGGTCAACACCGTCTCTATATTGCCACCAAT 918
Qy 261 ThrGlyGlyProTyrAspGlySerSerGlyAspValTrpLysPheSerValThrSerGly 280
Db 919 ACGGTGGTCCGTATACGGGAGCTCCGGCGACGCTCGGAAATTCCTCGTGACCTCCGG 978
Qy 281 ThrTrpThrArgIleSerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyr 300
Db 979 ACATGACCGCAATCAGCCCGGTACCTTCGACGACACCGGCAACGACTACTTTGGTTAC 1038
Qy 301 SerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSer 320
Db 1039 AGCGGCTCTCATTCGACCGCGACCCGCAACAGCATATGTTGGCAACCCAGATATCG 1098
Qy 321 TrpTrpProAspThrIleIlePheArgSerThrAspGlyGlyAlaThrTrpThrArgIle 340
Db 1099 TGGTGGCGCGACACCAATAATCTTTCCGAGCACCGCGCGGTGCGAGTGGACGCGATC 1158
Qy 341 TrpAspTrpTrpSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGlu 360
Db 1159 TGGGATGGACGAGTTATCCCAATCGAAGCTTCGGATATGCTTGATTCATTTCCGGCGAG 1218
Qy 361 ProTrpLeuThrPheGlyValGlnProAsnProValProSerProLysLeuGlyTrp 380
Db 1219 CTTTGGCTGACCTTCGGCGTACAGCCGATCTCCCGTACCCAGTCCGAGCTCGGCTGG 1278
Qy 381 MetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThrGly 400
Db 1279 ATGGATGAAGCGATGCGCAATCGATCGTTCAACTCTGTATCGGATGCTCTACGGAACAGGC 1338
Qy 401 AlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlyGlnIleHisIle 420
Db 1339 GCGAGCTTGTACGCAACAAATGATCTCAGCAAGTGGGACTCCGGCGCGCAGATTTCATATC 1398
Qy 421 AlaProMetValLysGlyLeuGluThrAlaValAsnAspLeuIleSerProProSer 440
Db 1399 GCGCGATGTCGCAAGGATGGAGGAGCGCGGTAACGATCTCATCAGCCCCCGCT 1458
Qy 441 GlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyPheThrHisAlaAspValThr 460
Db 1459 GCGCGCCGCTCATCAGCGCTCTCGGAGACCTCGCGGCTTCACCCAGCGCGAGTTACT 1518
Qy 461 AlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSerValAspTyr 480
Db 1519 GCGGTGCCATCAGCATCTTCAGTCACCGGTTCACGACCGCGCACCGGTCCGACTAT 1578
Qy 481 AlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspProSerSerClnPro 500
Db 1579 GCGGAATTAATCCGTCGATCATCGTTCGCGCTGGAAATTTTCGATCCATCAGGCAACCG 1638
Qy 501 AsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGlySerGlu 520
Db 1639 AACGACAGCAGCTCGCGTTCTCGACAGACGGCGGCAAGNACTGGTTTCCAAAGGACGGA 1698

521 ProGlyGlyValThrThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheVal 540
1699 CTGGCGGGGTACGACGACGGCGCGCATCGCTCGCGCATCGCGCGCATCGCTCTGTTTCGTC 1758
541 TrpAlaProGlyAspProGlyGlnProValValThrValAlaValGlyPheGlyAsnSerTrp 560
1759 TGGGCTCCCGCGCATCCCGGTACGGTGTGGTGTACGCGATCGGATTTGGCAACTCTCTGG 1818
561 AlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLys 580
1819 GTTGCTTCGAAGGTGTTCCCGCCATGCCAGATCGCTCAGACCGGGTGATCCAAAG 1878
581 ThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThrPhe 600
1879 ACTTTCTATGCCCTATCCCAATGAACCTTCTATCGAAGCACGACGCGCGGTGACATTC 1938
601 GlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHisAlaVal 620
1939 CAACCGGTCGCGCGCGGTCTTCCGAGCAGCGGTGCCGTGCTCATGTTTCCACGCGGTG 1998
621 ProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsn 640
1999 CTGGAAAGAGCGCATCTGTGGCTCGCTGCTCATCGAGCGGCTTTACCACTCAACCAAT 2058
641 GlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlaValAlaValGlyPheGly 660
2059 GCGCGCAGCAGTTGGTCTGCAATCACCGCGGTATCTCCGCGGTGAACGGGTATTTGGT 2118
661 LysSerAlaProGlySerSerTyrProAlaValPheValValGlyThrIleGlyGlyVal 680
2119 AAGTCTCGCCCGCGGTGTCATACCCAGCGCGTCTTTGTGTCGCGCACGATCGGAGCGTT 2178
681 ThrGlyAlaTyrArgSerAspCysGlyThrThrTrpValLeuIleAsnAspAspGln 700
2179 ACGGGCGGTACCGCTCCGACGACTGTGGACGACCTGGGTACTGATCAATGATGACCAG 2238
701 HisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgVal 720
2239 CACCAATACGGAATTTGGGACAAAGCAATCACCGGTGACCACCGAATTTACGCGGGTG 2298
721 TyrIleGlyThrAsnGlyArgGlyIleValTyrGlyAspIleGlyAlaProSerGly 740
2299 TACATAGGCACGAACGCGCGTGAATTTGATACGGGACATTTGGTGTGCGCGTCCGGA 2358

RESULT 3

US-10-156-761-1845
; Sequence 1845, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1845
; LENGTH: 2646
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2646)

US-10-156-761-1845
Alignment Scores:
Pred. No.: 5,87e-224 Length: 2646
Score: 2429,50 Matches: 438
Percent Similarity: 74.0% Conservative: 111
Best Local Similarity: 59.0% Mismatches: 174
Query Match: 60.2% Indels: 19
DB: 6 Gaps: 8
US-09-917-376-3 (1-740) x US-10-156-761-1845 (1-2646)
Qy 6 TyrThrTrpSerAsnValAlaIleGlyGlyGlyPheValAspGlyIleValPheAsn 25
Db 115 TACAGCTGGAAGAACGCCGCTCGACCGCGCGCTTCGTCGCCGGCATCGCTTCAAC 174
Qy 26 GluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrpAsp 45
Db 175 CGCTCCGAGAAACCTCGCTACGCCGCCACCGACATCGCGCGCGCTTACCGCTGGGCC 234
Qy 46 AlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyr 65
Db 235 GAGTCTCGAAGACTCGACCGCGCTGCTCGACTCGCTCGGCTGGAGGACCTGGGGGCAC 294
Qy 66 AsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaVal 85
Db 295 ACGGCTGTGAGCCCTCGCTCGACTCGCTCGACCCGACCAAGGTGTACGCGCGCGTC 354
Qy 86 GlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAspGln 105
Db 355 GGCACTGTACGAAACAGCTGGGACCGCGGCAACCGTGCCTGCTCAGGTCCGCGCACCG 414
Qy 106 GlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArg 125
Db 415 GCGCGGAGCTGGCAAGACCGACCTGCCCTTCAAGCTGGCGGGGAAATGCGCGCGCG 474
Qy 126 GlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAla 145
Db 475 GGCATGGCGGAGCGCTCGCGTGCACCCGAAACAGGACAGCGGTCTGTATCTCGCGCG 534
Qy 146 ProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThr 165
Db 535 CCCAGCGCAAGGGGCTGTGCGGTTCGACGACCTCGGGGCTCTCTGTCGCGAGGTACC 594
Qy 166 AsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGlnSer 185
Db 595 GACTTCCCGAACGTCGGCACCTACGTGACGACGCGCACGACACGCGGGGTACGCGTCC 654
Qy 186 AspIleGlnGlyValValTrpValAlaPheAspLysSerSerSerSerLeuGlyGlnAla 205
Db 655 GACAAACAGCGCATCTGTGGTCACTTCACGAGTGCACGGGTTCGCCGGGAGCTCC 714
Qy 206 SerLysThrIlePheValGlyValAlaAspProAsnAsnProValPheTrpSerArgAsp 225
Db 715 ACGCGGACGGTGTACGTGGGTTCGCGCAAGGACAACTCCGTCTATCGCTCCACGGAC 774
Qy 226 GlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPheIleProHisLysGly 245
Db 775 GCGGCGCGACCTGTGTCGCGCGCGCACCGCCCATCTCGCCGCCAACAGGCGC 834
Qy 246 ValPheAspProValAsnHisValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyr 265
Db 835 GTGCTGACCGCGGCAACCGGTGTCTGTACTTCGCTACAGGTACAGGCGCGGACCGGTAC 894
Qy 266 AspGlySerSerGlyAspValTrpLysPheSerValThrSerGlyThrTrpThrArgIle 285
Db 895 GACGCGCGCAAGGACAGCTGTGGGTACACGACGACGACCGGACCTTGGACGACATC 954
Qy 286 SerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyrSerGlyLeuThrIle 305
Db 955 AGCCCGGTTCGGAGGCGGACACCC-----TACTACGGCTTCAGCGGGGTGACCGGTG 1005
Qy 306 AspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTrpTrpProAspThr 325

Db 1006 GACCGCAGCATCCGGGACGGTGTACGACTCGGTACAGCTCTCTGGTGGCCGACACG 1065
 Qy 326 IleIlePheArgSerThrAspGlyGlyAlaThrTrpThrArgIleTrpAspTrpThrSer 345
 Db 1066 CAGCTCTTCGCTCCACGACAGCGGGCGGACCTTGGACGAAGGCGTGGACTACACCTCG 1125
 Qy 346 TyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGluProTrpLeuThrPhe 365
 Db 1126 TATCCGAGCCGCTCGAACCGCTTCCACATGATGTCTCGTCTCGCCCTGGCTCACCTCG 1185
 Qy 366 GlyValGlnProAsnProValProSerProLysLeuGlyTrpMetAspGluAlaMet 385
 Db 1186 GGACGGAACCCCGCACCGCCCGACGACAGCCCGAACTCGGCTGGATGACCGAGTCCCTG 1245
 Qy 386 AlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAla 405
 Db 1246 GAGATCGACCGCTTCGACTCGCGGCGATGATGACGAAGCGGGCGGCGTCTACGGC 1305
 Qy 406 ThrAsnAspLeuThrLysTrpAspSerGlyGlyGlnIleHisIleAlaProMetValLys 425
 Db 1306 ACGGACAACTTGCAGAACTGGGACAGCGGAGCCAGTTCACCATCAAGCCGATGGCGGG 1365
 Qy 426 GlyLeuGluThrAlaValAsnAspLeuIleSerProProSer---GlyAlaProLeu 444
 Db 1366 GGCCTGGAGGAGACGGCGCTCAACGACCTCGCTCGCTCCGCGGCGGCGCCAGCTG 1425
 Qy 445 IleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThrAlaValProSer 464
 Db 1426 TTCAGCGGCTCGGTGATCATCGCGGCTTCGGGACAGGACCTCACCGCGTCCGCTCG 1485
 Qy 465 ThrIlePheThrSerProValPheThrThrGlyThrSerValAspTyrAlaGluLeuAsn 484
 Db 1486 CTGATGTACAGCTCGCGCACTTCACACGAGCAGCAGCTCGACTAGCCGCGGACCGAC 1545
 Qy 485 ProSerIleIleValArgAlaGlySerPheAspProSerSerGlnProAsnAspArgHis 504
 Db 1546 CCGGCGCAGGTGGTGGCGGTTCGGAATCTCGAC-----TCGGGTCCG-----CAT 1590
 Qy 505 ValAlaPheSerThrAspGlyGlyLeuAsnTrpPheGlnGlySerGluProGlyGlyVal 524
 Db 1591 GTGGCGTTCGACGGACACCGCGCACTGGTTCCGGGGCGGACCTTCGGGGGTC 1650
 Qy 525 ThrThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaProGly 544
 Db 1651 AGCGGGGTGGAGCGTTCGGCGCGCTCCGACGCGAGTCTGCTGTGTGGAGCGCGCG 1710
 Qy 545 AspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGln 564
 Db 1711 GGCACCGGG-----GTGCGAGTACACGACCGGGTTCCGACCTCGTGTGGGTCCGCG 1764
 Qy 565 GlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThrPheTyrAla 584
 Db 1765 GGCCTCCGCGCGGGGGGATGTCGAGTCCGACCGGGTTCGACCGGAGACCTTCTACGGC 1824
 Qy 585 LeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThrPheGlnProValAla 604
 Db 1825 TTCAGTCCGCGAGTTCAGTTCAGTTCGACCGCGGGCGGACCTTCACGGGTCCGCG 1884
 Qy 605 Ala---GlyLeuProSerSerGlyAlaValGlyValMetPheHisAlaValProGlyLys 623
 Db 1885 GCCACGGGCTCCGAGCGGCGACAGC-----GTGCGCTTCAAGGCGCTCCCGGCGACG 1938
 Qy 624 GluGlyAspLeuTrpLeuAlaAlaSerSer-----GlyLeuTyrHisSer 638
 Db 1939 AAGGCGCACATCTGGCTGGCGGGCGGCGGAGCGGCGGTACCGGCTGTGGCACTCG 1998
 Qy 639 ThrAsnGlySerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnValGly 658
 Db 1999 ACGACGCGCGCGCGGCTTCCACCAAGCTCGCCACCGCTCGACCGGCGGACCATCGGC 2058
 Qy 659 PheGlyLysSerAlaProGlySerSerTyrProAlaValPheValValGlyThrIleGly 678

Db 2059 TTCGCAAGGCGGCGACCGCGCTCGTACACAGCGCTCTACACGACGCGGAATCGGC 2118
 Qy 679 GlyValThrGlyAlaTyrArgSerAspCysGlyThrTrpValLeuIleAsnAsp 698
 Db 2119 GGTGTGCGCGCATCTTCGGTCCGACCGACAAAGGCGCGAGCTGGACCGCGTCAACGAC 2178
 Qy 699 AspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArg 718
 Db 2179 GATGCCCACTAGTGGGTGGAGCGGCGCGCGATCACCGGTGACCCCGAGGTCTACGGG 2238
 Qy 719 ArgValTyrIleGlyThrAsnGlyArgGlyIleValTyrGlyAspIleGlyGlyAlaPro 738
 Db 2239 CGCGTGTATGTCTCGACGAACCGGCGCGGATCTCTACGCGGACACCGCGGCTCTTCG 2298
 Qy 739 SerGly 740
 Db 2299 GACGGC 2304
 RESULT 4
 US-10-156-761-1
 ; Sequence 1, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 1
 ; LENGTH: 9025608
 ; TYPE: DNA
 ; ORGANISM: Streptomyces avermitilis
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (4187715)
 ; OTHER INFORMATION: a, t, c, g, other or unknown
 US-10-156-761-1
 Alignment Scores:
 Pred. No.: 1,86e-219 Length: 9025608
 Score: 2429.50 Matches: 438
 Percent Similarity: 74.0% Conservative: 111
 Best Local Similarity: 59.0% Mismatches: 174
 Query Match: 60.2% Indels: 19
 DB: 6 Gaps: 8
 US-09-917-376-3 (1-740) x US-10-156-761-1 (1-9025608)
 Qy 6 TyrThrTrpSerAsnValAlaIleGlyGlyGlyPheValAspGlyIleValPheAsn 25
 Db 2277508 TACAGTGGAGAGACCGCGCGTCCAGCGCGGCGTCTCGTCCCGGACATCGTCTCAAC 2277567
 Qy 26 GluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrpAsp 45
 Db 2277568 CGCTCCGAGAGAAACCTCGCTACCGCCGACCGGCGGCGCTTACCGCTGGGCGC 2277627
 Qy 46 AlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyr 65
 Db 2277628 GAGTCTCTGAAGACCTCGACCGCGCTGTCTCGACTCGGTGGAGCGACCTGGGGGCGAC 2277687
 Qy 66 AsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaVal 85

Db 2277688 ACGGGTGTGAGCCTCGCTCGACTCCGTGACCCGGAACAGGTGTACCGCGCGCTC 2277747
Qy 86 GlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAspGln 105
Db 2277748 GGCACGTACACGACAGCTGGGCAACCGGTGCTCGGTCTCAGGTCCGGGACCGG 2277807
Qy 106 GlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArg 125
Db 2277808 GGGCGAGCTGGGAGAGACGACCTGCCCTTCAAGCTGGGCGGAACATGCGGGCGG 2277867
Qy 126 GlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAla 145
Db 2277868 GGCATGGCGGAGCGGCTCGGGTCGACCCGAAACAGGACAGCGTGTGTATCTCGGCGG 2277927
Qy 146 ProSerGlyLysLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThr 165
Db 2277928 CCCAGCGCAAGGGGCTGTGGGTGACGAGACTCGGGGGCTCTCGTGCAGGTCAAC 2277987
Qy 166 AsnPheProaspValGlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGlnSer 185
Db 2277988 GACTTCCCGAACGTCGGCACCTACGTGACGAGCAGCGACCGACAGCGGGTACGCGTCC 2278047
Qy 186 AspIleGlnGlyValValTrpValAlaPheAspLysSerSerSerSerSerLeuGlyGlnAla 205
Db 2278048 GACACACAGGGCATCTGTGGTCACTTCAGAGTCGACGGGTCCCGGGGAGCTCC 2278107
Qy 206 SerLysThrIlePheValGlyValAlaAspProAsnAsnProValPheTrpSerArgAsp 225
Db 2278108 ACGCGGACGGTGTACGTGGGGTCCGCGACAAAGGACAACTCCGTCTATCGTCCACGGAC 2278167
Qy 226 GlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPheIleProHisLysGly 245
Db 2278168 GCGGGCGCGACCTGGTCCGCTGGCGCGCAGCCACCGGCCATCTCGCCACAAAGGCG 2278227
Qy 246 ValPheaspProValAsnHisValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyr 265
Db 2278228 GTGCTGACCGCGGAACCGGTGTGTACTCTCGCTACGAGCAGACAGGCGGACCGGTAC 2278287
Qy 266 AspGlySerSerGlyAspValTrpLysPheSerValThrSerGlyThrTrpThrArgile 285
Db 2278288 GACGGCGCAAGGACAGCTGTGGGTACACGACGAGACCGGGACCTGGACGAACATC 2278347
Qy 286 SerProValProSerThrAspThrAlaAsnAspTyrPheGlyTyrSerGlyLeuThrile 305
Db 2278348 AGCCCGGTCCGAGGCGGACAC-----TACTACGGCTTCAGCGGGCTGACCGGTG 2278398
Qy 306 AspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTrpTrpProAspThr 325
Db 2278399 GACCGGCAGCATCCGGGCGGACGGTGATGGCGACTGCGTACAGCTCCTGGTGGCGGACAG 2278458
Qy 326 IleIlePheArgSerThrAspGlyGlyAlaThrTrpThrArgileTrpAspTrpThrSer 345
Db 2278459 CAGCTCTTCGCTCCACGACAGCGCGGACCTGGACGAGGCTGGGACTACACCTCG 2278518
Qy 346 TyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGluProTrpLeuThrPhe 365
Db 2278519 TATCCGAGCGCTCGAACCGCTTCACCATGATGTCTCGTCTCCCTCGCTCGCTCAGCTGG 2278578
Qy 366 GlyValGlnProAsnProValProSerProLysLeuGlyTrpMetAspGluAlaMet 385
Db 2278579 GGAGCGAACCCCGCGCGCGAGACACCCGAAACTCGGTGATGACCGAGTCCCTG 2278638
Qy 386 AlaIleaspProPheAsnSerAspArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAla 405
Db 2278639 GAGATGACCCGTTCGACTCCGCGCGCATGATGTACGGAACGCGCGGCTGTACGGC 2278698
Qy 406 ThrAsnAspLeuThrLysTrpAspSerGlyGlyGlnIleHisIleAlaProMetValLys 425
Db 2278699 ACGGACAACTCGACAACTGGGACGCGGAGCCAGTTCACATCAGCCGATGGCGCGG 2278758
Qy 426 GlyLeuGluThrAlaValAsnAspLeuIleSerProProSer---GlyAlaProLeu 444
Db 2278759 GGCCTGGAGGAGACGGCGGTCAACGACCTCGCTCGCTCCCTCCGCGCGGCCCGACCTG 2278818

Qy 445 IleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThrAlaValProSer 464
Db 2278819 TTCAGCGCGCTCGGTGACATCGCGGCTTCGGGCACACGGACCTCACACGGTCCGCTCG 2278878
Qy 465 ThrIlePheThrSerProValPheThrThrGlyThrSerValAspTyrAlaGluLeuAsn 484
Db 2278879 CTGATGTACAGTCCCGAACTTCACACGAGACACCACTTCGACTACCGCAGACCGAC 2278938
Qy 485 ProSerIleIleValArgAlaGlySerPheAspProSerSerGlnProAsnAspArgHis 504
Db 2278939 CCGGGCAGGTGGTGGGTCCGCAATCTCGAC-----TCGGGTCCG-----CAT 2278983
Qy 505 ValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGlyVal 524
Db 2278984 GTGGCTTCTCGACGGACAAACCGGCCCAACTGGTTCGGCGGGCGGACCCCTTCGGGGGTC 2279043
Qy 525 ThrThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaProGly 544
Db 2279044 AGCGGGGTGGGACGGTCCGCGCGCGCTCCGACGGCAGTCTCGTGTGTGGAGCCCGCG 2279103
Qy 545 AspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGln 564
Db 2279104 GGCACCGGG-----GTCCAGTACAGACCGGGTTCGGCACCTCTGTGTGGGTCCCGG 2279157
Qy 565 GlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThrPheTyrAla 584
Db 2279158 GGCCTCCCGCGCGGGCGATCGTCCAGTCCGACCGGTGCGACCGGACCTTCACGGC 2279217
Qy 585 LeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThrPheGlnProValAla 604
Db 2279218 TTCAAGTCCGCGAGGTCTACGTCTAGTTCGACGCGCGGGGAGACCTTCACGGCTCCGG 2279277
Qy 605 Ala---GlyLeuProSerSerGlyAlaValGlyValMetPheHisAlaValProGlyLys 623
Db 2279278 GCCACGGGCTTCGCGAGCGCGCAGC-----GTGCGCTTCAAGCGCTGCCCGGACG 2279331
Qy 624 GluGlyAspLeuTrpLeuAlaAlaSerSer-----GlyLeuTyrHisSer 638
Db 2279332 AAGGGCAGCATCTGGCTGGCGGGCGCGGCGGCGGCGGCTAGCGGCTGTGGCACTCG 2279391
Qy 639 ThrAsnGlyLysSerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnValGly 658
Db 2279392 ACGGACGGCGCGCGGCTTCACCAAGCTCCGACCGTCCGACCGGCGGACACCATCGC 2279451
Qy 659 PheGlyLysSerAlaProGlySerSerTyrProAlaValPheValGlyThrIleGly 678
Db 2279452 TTCGCAAGCGCGGCGGCGGCGGCTCGTACGAGCGCTCTACACCGCGCGGAGATCGC 2279511
Qy 679 GlyValThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsnAsp 698
Db 2279512 GGTGTGGCGGCATCTTCGGGTCCGACCGCAAGGCGCGGAGCTGGACCGCGCTCACGAC 2279571
Qy 699 AspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArg 718
Db 2279572 GATCCACACAGTGGGTGGACGGCGCGGATCACCCTGACCGGTGACCCAGGCTTACGGG 2279631
Qy 719 ArgValTyrIleGlyThrAsnGlyArgGlyIleValTyrGlyAspIleGlyAlaPro 738
Db 2279632 CGCGTGTATGTGTGCGAAGCGGCGGGATCGTACGCGGACACCGCGGCTCTTCG 2279691
Qy 739 SerGly 740
Db 2279692 GACGGC 2279697
RESULT 5
US-10-026-994-4
; Sequence 4, Application US/10026994
; Publication No. US20030113732A1
; GENERAL INFORMATION:
; APPLICANT: Dunn-Coleman, Nigel
; APPLICANT: Goedegebuur, Frits
; APPLICANT: Ward, Michael

```

; APPLICANT: Yao, Jian
; TITLE OF INVENTION: EGY Endoglucanase and Nucleic Acids
; FILE OF INVENTION: Encoding the Same
; FILE REFERENCE: GC698
; CURRENT APPLICATION NUMBER: US/10/026,994
; CURRENT FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 2517
; TYPE: DNA
; ORGANISM: Trichoderma reesei
US-10-026-994-4

Alignment Scores:
Pred. No.: 7,51e-142 Length: 2517
Score: 1579.00 Matches: 325
Percent Similarity: 59.3% Conservative: 123
Best Local Similarity: 43.0% Mismatches: 264
Query Match: 39.1% Indels: 44
DB: 5 Gaps: 20

US-09-917-376-3 (1-740) x US-10-026-994-4 (1-2517)

QY 6 TyrThrTrpSerAsnValAlaIle---GlyGlyGlyGlyPheValAspGlyIleValPhe 24
Db 61 TTTTTCATGGAAGAACGTCAGCTCGCGCGCGCGGCTTCGTCCCGGCATCATCTTC 120
QY 25 AsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrp 44
Db 121 CATCCCAAGACAAAGCGGTAGCATATATGCAACAGATATTGGCGGCTGTACCGCCTC 180
QY 45 AspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpVal-----GlyTrp 60
Db 181 AAC---GCGCAGCACTCATGACCGCGCTCAGGATGGGATTCGTATATGCCGCTGG 237
QY 61 AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLys 80
Db 238 CACAACCTGG-----GGCATCGACGCTGTTGCGCTTCATCCGAGGACGATCAAAAG 288
QY 81 ValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyIleAlaLeu 100
Db 289 GTGTATGCCGAGTCGGCATGATATACGAACAGCTGGGATCCGAGTAATGGAGCCATCAT 348
QY 101 ArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGly 120
Db 349 CGCTCGTCAGACCGCGCGCCACACGCTGCTCTCACCACTTGCCCTTCAAGATCGGGGT 408
QY 121 AsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIle 140
Db 409 AACATGCCAGACGCGGAGCGGAGAGCGCTGCGTGTGATCCGCGCAACTCCACATC 468
QY 141 LeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThr 160
Db 469 ATCTACTATTGGTGTCTGCTCAGGAACCGGCTCTCGGAAGTCTACGACGCGCGCGTGACC 528
QY 161 TrpSerGlnMetThrAsnAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThr 180
Db 529 TTTTCCAAAGTCTCGTGGTTCACGGCAACTGGGACGTACATCCACCGCGAGTGAATCT 588
QY 181 ThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAspLysSerSer 200
Db 589 AACGGCTACAAACGCGACAAAGCAAGGACTCATGTGGGTTACGTTTCGATTCACACGACG 648
QY 201 SerLeuGlnAlaSerLysThrIlePhePheValGlyValAlaAspPro---AsnAsnPro 219
Db 649 ACGACCGGGGAGCGACGCTCTCGTATCTTTGTGGCAGCGCTCATACATCATCTGCTTCA 708
QY 220 ValPheTrpSerArgAspGlyGlyValThrTrpGlnAlaValProGlyAlaProThrGly 239
Db 709 GTCTATGTGACACGAATCGCGGCTCCACGTGGAGTGTGTACCGGGGAGCCAGGGGAAA 768
QY 240 PheIleProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSer 259

```

1816	GT	CAGCAG	AAGACACCGC	CAGCAGCTTC-----ACGCGCGGG---CCCAAGCTGGGC	1863															
612	Ala	Val	Gly	ValMet-----PheHis	Ala	Val	Pro	Gly	Leu	Gly	Asp	Leu	Trp	Leu	629					
1864	AGC	G	CAGG	AGAGATCGG	GATATCGT	CTACCCG	ACACCGCGG	GCGACGTTGT	ATGTC	1923										
630	Ala	Ala	Ser	Ser	Gly	Leu	Thr	His	Thr	Asn	Gly	Gly	Ser	Ser	Trp	Ser	Ala	Ile	---648	
1924	TGC	ACC	CGC	AGTCCG	CATATTCGCT	CCAC	AGACTCGG	CA	AGACCTTGG	CCCAAGTCTCC	1983									
649	Thr	Gly	Val	Ser	Ser	Ala	Val	Asn	Val	Gly	Phe	Gly	Leu	Ser	Ala	Pro	Gly	Ser	Ser	Tyr668
1984	ACG	CCCT	GACCA	CAACCTTAC	CAGATCG	CCCTG	GGTCTGG	GCTCA---GGCTCGA	ACTGG	2040										
669	Pro	Ala	Val	Phe	Val	Val	Gly	Thr	Ile	Gly	Gly	Val	Thr	Gly	Ala	-----Tyr	Arg	Ser	686	
2041	---AAC	CTGTAT	GCCTTC	CGCACC-----GGCCCGT	CAGGGGCTCG	CTCTAC	GCACGT	2091												
687	Asp	Asp	Cys	Gly	Thr	Thr	Trp	Val	Leu	Ile	Asn	Asp	Gln	His	Gln	Thr	Gly	Asn	Trp706	
2092	GG	AG	CAG	CGCGGCCT	CTCTG	AGCGG	ACATCCAG	GGCTCC	CAGGGCTCG	GCTCCAT	CGC	AC	2151							
707	Gly	Gln	Ala	Ile	Thr	Gly	Asp	His	Ala	Asn	Leu	Arg	Val	Tyr	Ile	Gly	Thr	Asn	Gly726	
2152	AG	CAC	CAAGTCC	CGCGCAG	CAGCAGCCG	CGGCGA	AGTCTAC	TG	GGCGACCA	CGGC	2211									
727	Arg	Gly	Ile	Val	Tyr-----Gly	Asp	Ile	Gly	Val	Ala	Pro	Ser	Gly	740						
2212	CGG	GGG	CGCTTTT	TACGCTCAG	GAACCGT	CGGGCGG	CGCAG	CGGGCGG	2259											

Qy	61	AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAenThrAsnLys	80
Db	278	CAAACTGG-----GGCATCGACCGCTGTTGCGCTTCATCGGAGGACGATCAAAAG	328
Qy	81	ValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeu	100
Db	329	GTGTATCCCGAGTCGGCATGTATACGAAACAGCTGGGATCCGAGTAAATGAGGCATCAT	388
Qy	101	ArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGly	120
Db	389	CGCTCGTCAGACCGCGCGCAACGTGGTCTTCACCAACTTGGCCCTTCAAGTCGGGGT	448
Qy	121	AsnMetProGlyArgGlyMetGlyAluArgLeuAlaValAspProAsnAspAsnIle	140
Db	449	AACATGCAGGACGGGAGCGGAGAGCGTCTGGCTCGATCCGGCCAACTCCAAATC	508
Qy	141	LeuTyrPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThr	160
Db	509	ATTCTACTTTTGGTGTCTCGCTCAGGAAACGGGCTCTGGAAGTCTACGAGCGCGCTGACC	568
Qy	161	TrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThr	180
Db	569	TTTTCCAAAGTCTCGTGTTCACCGCCAACTGGGAGCTATCCTCCAGACCCGAGTATTC	628
Qy	181	ThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAspLysSerSer	200
Db	629	AACGGCTACAAACAGCGACAAAGGAACTCATGTGGTTACGTTTCGACTCAACACGACGC	688
Qy	201	SerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspPro---AsnAsnPro	219
Db	689	ACGACCGGGGAGCCACGTCGTATCTTTGTTGGACCGGCTGATAACATCACTGCTTCA	748
Qy	220	ValPheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGly	239
Db	749	GTCTATGTGACAGCAATGCGCGCTCCAGCTGAGTGTGTACCGGGAGCCAGGGAAA	808
Qy	240	PheIleProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAlaThrSer	259
Db	809	TACTTTCTTCAAGGCGAACTGCACCCAGCAGACGAGGCGCTTGTATCTCGACTATTC	868
Qy	260	AsnThrGlyGlyProTyrAspGlySerSerGlyAspValTrpLysPheSerValThrSer	279
Db	869	GATGGCACAGGGCGGTATGATGTCACACTTGGCTCAGTGTGGAGGTACGACATTCGACGG	928
Qy	280	GlyThrTrpThrArgIleSerProValProSerThrAspThrAlaAsnAspTyrPheGly	299
Db	929	GGAACTTGGAAAGACATCACCCCTGTCTCTGGAACTCAGATCTA-----TACTTTGGC	979
Qy	300	TyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIle	319
Db	980	TTTGGCGGCTTGGCCCTCGATTGTCGAAAGCCAGGAAACCTTGTGTCTTCTTTGAAC	1039
Qy	320	SerTrpTrpProAspThrIleIlePheArgSerThrAspGlyAlaThrTrpThrArg	339
Db	1040	TCCTTGTGGCCAGATGCTCAGCTGTTTGGTTCAGCAGCTCTGGGACAAACATGGAGCCCG	1099
Qy	340	IleTrpAspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAla	359
Db	1100	ATCTGGCGGTGGCGAGCTATCCGACTGAGACCTATTACTACAGCATCTCAACTCCCAA	1159
Qy	360	GluProTrpLeuThrPheGly---ValGlnProAsnProValProSer-----	375
Db	1160	GCACCTGGGATCAAGAACAACTTTATTCGATGTGACGAGCGAGTCCCGTCGATGCTCTC	1219
Qy	376	---ProLysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArg	394
Db	1220	ATCAAGCGCTCGCTCGATGATTGATGCTCTCGAGATTGACCCCAACCGACACCAACC	1279
Qy	395	MetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSer	414
Db	1280	TGGCTCTACCGCACCGGAATGACAACTTTTGGCGGCGCAGATCTCAACAACTCGGACAG	1339

QY 415 GlyGlyGlnIleHisIleAlaProMetValIysGlyLeuGluGluThrAlaValAsnAsp 434
 Db 1340 CCGCATGTGTCAATCAATCACTGCAGACGGCATCGAGAAATTCCTCCGTCAGGAC 1399
 QY 435 LeuIleSerProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPhe 454
 Db 1400 CTGGCTCTGCACCGCGGAGAGAGCTATTGGCCGAGTCTGGAGACGACAAAGCTTC 1459
 QY 455 ThrHisAla-----AspValThrAlaValProSerThrIlePheThrSerProVal 471
 Db 1460 ACCTTGGCAGCAGAAACGACCTCGGACATCGCGCAGACGCTTGGGCAACGCCACA 1519
 QY 472 PheThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAla 491
 Db 1520 TGGGCGACCTCGACGAGCGTGCACCTCGCGGAACTCGGTCAAGAGCGTCTGTCGCGTC 1579
 QY 492 GlySerPheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGly 511
 Db 1580 GGCACACCGCGCGCACGACAG-----GTGGCATCTCGTCCGACGCG 1624
 QY 512 GlyLysAsnTrpPheGlnGlySerGluProGlyGlyValThrThrGlyGlyThrValAla 531
 Db 1625 GCGCGACGTGGAGCATCGACTACGCGCGCACAGCTCATGACGCGGCGACGTGGCC 1684
 QY 532 AlaSerAlaAspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProValVal 551
 Db 1685 TATTCGCGCGAGCGGACACGACATCTCTGTGACCGCTCGTCCGCG-----GTG 1735
 QY 552 TyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGln 571
 Db 1736 CAGCGCTCGAGTTCACGCGGAGCTTTCTGCTCGCTCGAGCTCGCGCGCGCGCGCTC 1795
 QY 572 IleArgSerAspArgValAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyr 591
 Db 1796 ATCGCTCGGACAGAACCAACAGCGCTCTCTACGCGCGCTCGGATCGACCTTTTAC 1855
 QY 592 ArgSerThrAspGlyGlyValThrPheGlnProValAlaAlaGlyLeuProSerSerGly 611
 Db 1856 GTCAGCAAGGACACCGGACGAGCTTC-----ACGCGCGG---CCCAAGCTGGGC 1903
 QY 612 AlaValGlyValMet-----PheHisAlaValProGlyLysGluGlyAspLeuTrpLeu 629
 Db 1904 AGCGCAGGAGCATCCGGATATCGCTGCTCATCCCGACACCGCGGCGACGTTGTATGTC 1963
 QY 630 AlaAlaSerSerGlyLeuTyrHisSerThrAsnGlyGlySerSerTrpSerAlaIle--- 648
 Db 1964 TGCACCGAGCTCGGATATTCGCTCCACAGCTCGGCGACACCTTTGGCCAACTCTCC 2023
 QY 649 ThrGlyValSerSerAlaValAsnValGlyPheGlyLysSerAlaProGlySerSerTyr 668
 Db 2024 ACCGCGCTGACCAACACCTACCATCGCTCGCTGGTGTGGGTCA---GGCTCGAACTGG 2080
 QY 669 ProAlaValPheValGlyThrIleGlyValThrGlyAla-----TyrArgSer 686
 Db 2081 ---AACTGTATGCTTCGGACAC-----GGCCGCTCAGGGGCTCGCTCTACGCCAGT 2131
 QY 687 AspAspCysGlyThrThrTrpValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnTrp 706
 Db 2132 GGAGACGCGCGCTCTCTGACGAGACATCCAGGCTCCCGGCTCGCTCCATCCATCCAG 2191
 QY 707 GlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgArgValTyrIleGlyThrAsnGly 726
 Db 2192 AGCACCAGGTGCGCGGAGCGGAGCAGCCGCGGCGAAGTCTACGTGGGACCAACGCGC 2251
 QY 727 ArgGlyIleValTyr-----GlyAspIleGlyAlaProSerGly 740
 Db 2252 CGGGCGCTTTTACGCTCAGGGAACCTCGCGCGCGGACCGCGCGG 2299

RESULT 7

US-10-156-761-2561
 ; Sequence 2561, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 2561
 ; LENGTH: 2217
 ; TYPE: DNA
 ; ORGANISM: Streptomyces avermitilis
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(2217)
 ; US-10-156-761-2561

Alignment Scores:
 Pred. No.: 4,07e-125 Length: 2217
 Score: 1405.00 Matches: 308
 Percent Similarity: 54.6% Conservative: 104
 Best Local Similarity: 40.8% Mismatches: 257
 Query Match: 34.8% Indels: 86
 DB: 6 Gaps: 21

US-09-917-376-3 (1-740) x US-10-156-761-2561 (1-2217)

QY 6 TyrThrTrpSerAsnValAlaIleGlyGlyGlyPheValAspGlyIleValPheAsn 25
 Db 124 TACCGCTGCGGCAACGCGCTCATCGGGGACCGCGTCTCGTACCAGCGTCTTCCAC 183
 QY 26 GluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrpAsp 45
 Db 184 CCTCCGTACGCGGTCTCGCTACGCCGCGGACCGACATCGCGCGCGCTTACCGCTGGAC 243
 QY 46 AlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyr 65
 Db 244 GACCGCGCGCGCTGGACCGCTCATCGACACCTCGCTGGGACGCTGGAACCTC 303
 QY 66 AsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaVal 85
 Db 304 CTCGCGCTCGAGCGATGCGCTCGACCCCGACCGCGCGCGCTCTACCTCGCGTG 363
 QY 86 GlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAspGln 105
 Db 364 GGCACCTACGCCAGTCTGTTGGGCGGCAAC---GGCGGCTCTCGGCTCCGAGGACCGC 420
 QY 106 GlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArg 125
 Db 421 GCGCGCACCTGACCGCGCACCGCTGACCTGGAAGCTCGCGGCGCAACGAGGACGCGCGC 480
 QY 126 GlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyVala 145
 Db 481 GCGCGCGGTGAGCGACTCTCTGTCACCGCGCGGACGACACCTCTGGCTGGGCGACG 540
 QY 146 ProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThr 165
 Db 541 ---CGGCACGACGCGGTCTGCTCAAGTCGACCGCGCGGCGCGCCACTTGGGCGCGCGAC 597
 QY 166 AsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGlnSer 185
 Db 598 GCCTTCCCG-----CGGAAGGCGCAACTCC 621
 QY 186 AspIleGlnGlyValValTrpValAlaPheAspLysSerSerSerSerGlyGlnAla 205

Db 622 TCCGGCAGGAGTCGTTC-----CTCGTCGCCGC 654
Qy 206 SerLysThrIlePheValGlyValAlaAspProAsnAsnPro-----Val 220
Db 655 GGGCGCAGCGTCTACCGCGGTGGGTGACGCGCAGCGCCTCGGCGCGCGAACCTG 714
Qy 221 PheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPhe 240
Db 715 TACCGCAGCGCGAC--GGCAGCACCTGGGGCGGTTCCTCGCGCGCGCTCGCGCAC 771
Qy 241 -----IleProHisLysGlyValPheAspProValAsnHisValLeuTyrIleAla 257
Db 772 TCCGCCAGGTCCCGTCGCGCGGTGACACGCGCACCCCGAGCTGTACTGACG 831
Qy 258 ThrSerAsnThrGlyGlyProTyrAspGlySerGlyAspValTrpLysPheSerVal 277
Db 832 TACGGCAGCGCACCG 891
Qy 278 ThrSerGlyThrTrpThrArgIleSerProVal-----ProSerThrAspThr 293
Db 892 GCCACGGGAGCTGACCGAGGTCACTCCCGGTGAAGCGCGCGCGCGCGCGCGCG 951
Qy 294 AlaAsnAspTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIle 313
Db 952 TCGCGGACACCTTCGCTACCGCGGGTCCGCTCGACCGCGCGCGCGCGCGCGCG 1011
Qy 314 MetValAlaThrGlnIleSerTrpTrpProAspThrIleIlePheArgSerThrAspGly 333
Db 1012 GTCTCTCCACCAACACCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1071
Qy 334 GlyAlaThrTrpThrArgIleTrpAspTrpThrSerTyrProAsnArgSerLeuArgTyr 353
Db 1072 GCGCGTACCTGCGAGTCCCTCAAGGAC-----GCCGCC 1104
Qy 354 ValLeuAspIleSerAlaGluProTrpLeuThrPheGlyValGlnProAsnProVal 373
Db 1105 GTGTTCAGCGTCCGAGACTCCCTTCCTCGACTGGCGCGCAGC----- 1149
Qy 374 ProSerProLysLeuGlyTyrMetAspGluAlaMetAlaIleAspProPheAsnSerAsp 393
Db 1150 ---NAGCGGAGTTCGGCTGGTGTATCCAGCGCTCGCGGTGCGACCGCGTACGCTCCAG 1206
Qy 394 ArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAsp 413
Db 1207 CACGTCTGTACGGACCGCGCGCGACCTCTACGCGCACCGCGCGCTCAAGCGTGG-- 1263
Qy 414 SerGlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluThrAlaValAsn 433
Db 1264 -----GCACCGCGGATCCGCGCGCTGGAGGAGAGCGCGCGTGGCG 1302
Qy 434 AspLeuIleSerProSerGly---AlaProLeuIleSerAlaLeuGlyAspLeuGly 452
Db 1303 CAACTGATCTCGCCCGCGTCCGGAGGACACCTGATCAGCGACTCGGGGACATCGT 1362
Qy 453 GlyPheThrHisAlaAspValThrAlaValProSerThr-----IlePheThrSerProVal 471
Db 1363 GTGATGTACACGAGCGCTCACGCGCTCTCGTCCGCGCGCATGGCGCGCGCGCGGTG 1422
Qy 472 PheThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAla 491
Db 1423 TTCGGTCCGCGACGGGACTCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1482
Qy 492 GlySerPheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGly 511
Db 1483 GCGTGGCGCGCACCGCACCGC-----GCGTACTCCACGACGCGC 1524
Qy 512 GlyLysAsnTrp-----PheGlnLysSerGluProGlyGlyValThrThrGlyGlyThr 529
Db 1525 GGGGGGACCTGGGCGCGCTTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1584
Qy 530 ValAlaAlaSerAlaAspGlySerArgPheValTrpAla-----ProGlyAsp 545
Db 1585 ATCGCCACAGTGCAGCGCGCGCGCACACTGTGTGTCTTCTGTCGTCGTCGTCGTCG 1641

Qy 546 ProGlyGlnProValValTyrAla-----ValGlyPheGlyAsnSerTrpAla 561
Db 1642 -----ACGACGTACGCGCGCCACCGCTCGACGCAACACGCGCGAGCTGCTCC 1689
Qy 562 AlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThr 581
Db 1690 GAGGTCTCTCTCTCCCGAAGCGCGCACACCGGTGCGCGACCGCGCGCGCGCGCG 1749
Qy 582 PheTyrAla-----LeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThr 599
Db 1750 TTCTACGCTACGACTTCGACATGAACTATACGCGCAGCAGCTGACAGTGGCGCTCG 1809
Qy 600 PheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHisAla 619
Db 1810 TTCACGCGCGTTCGCGCGGCGACTGCGC---TCCGCGCAGCAGCAGTCAAGCTGCGCG 1866
Qy 620 ValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSer---SerGlyLeuTyrHisSer 638
Db 1867 GCGCGCGGACGAAGCGCGGACCTTGTGGCTCTCCGCAATGGAACGGGCTCTACCGGTCC 1926
Qy 639 ThrAsnGlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnValGly 658
Db 1927 ACCGACGCGGCGACCTTCCAGGATCGACAGTCTGGGCTCTGACACCTTCGCGC 1986
Qy 659 PheGlyLysSerAlaProGlySerSerTyrProAlaValPheValValGlyThrIleGly 678
Db 1987 TTCGCAAGCG 2046
Qy 679 GlyValThrGlyAlaTyrArgSerAspCysGlyThrThrTrpValLeuIleAsnAsp 698
Db 2047 ACCATCACCGCGCTACCGCTCCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2106
Qy 699 AspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArg 718
Db 2107 GACGCCACCACTGGGGGTGGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2166
Qy 719 ArgValTyrIleGlyThrAsnGlyValArgGlyIleValTyrGlyAsp 733
Db 2167 CGGTCTACCTCGCCACCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2211

RESULT 8

US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Alignment Scores:

Pred. No.:	1.61e-120	Length:	9025608
Score:	1405.00	Matches:	308
Percent Similarity:	54.8%	Conservative:	104
Best Local Similarity:	40.8%	Mismatches:	257
Query Match:	34.8%	Indels:	86
DB:	6	Gaps:	21
US-09-917-376-3 (1-740) x US-10-156-761-1 (1-9025608)			
Qy	6	TyrThrTrpSerAsnValAlaIleGlyGlyGlyPheValAspGlyIleValPheAsn	25
Db	3161882	TACCGTGGCGCAACCGCGTATCGGGGACCGGCTTCGTACCGGGCGTCTTCCAC	3161823
Qy	26	GluGlyAlaProGlyIleLeuTyrrValArgThrAspIleGlyMetTyrrArgTrpAsp	45
Db	3161822	CCCTCCGTACCGCGTCTCGCCTACGCCCGGACACATCGCGCGCGCTACCGTGGGAC	3161763
Qy	46	AlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyrr	65
Db	3161762	GACCGCGCGCGCGTGGACCCCGCTCATCGACACCTCGGCTGGGACGACTGGAACTC	3161703
Qy	66	AsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaVal	85
Db	3161702	CTCGCGTTCGAGGCGATGCGCTGACCCACGACCCCGGACCGGCTCTACTCGCCGTG	3161643
Qy	86	GlyMetTyrrAsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAspGln	105
Db	3161642	GGCACCTACGCCCGAGTCGTGGCGGGCAAC--GGCGGCGCTCTCGCTCCGAGACCGC	3161586
Qy	106	GlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArg	125
Db	3161585	GGCGCACCTGGACCCCGCACCGACTGACCGTGAAGCTCGCGGCAACGAGGACGGCGC	3161526
Qy	126	GlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrrPheGlyAla	145
Db	3161525	GCGCGGTGAGGCACTCTCTGACCCCGGACACGCAACCCCTCTGGCTGGGACG	3161466
Qy	146	ProSerGlyLysGlyLeuTyrrArgSerThrAspSerGlyAlaThrTrpSerGlnMetThr	165
Db	3161465	--CGGCACGACGGGTGCTCAAGTCGACCGACCGCGGCGCCACTTGGGACCGCGACC	3161409
Qy	166	AsnPheProAspValGlyThrTyrrIleAlaAsnProThrAspThrThrGlyTyrrGlnSer	185
Db	3161408	GCCTTCCCG-----GCGAAGCGCAACTCC	3161385
Qy	186	AspIleGlnGlyValValTrpValAlaPheAspLysSerSerSerSerLeuGlyGlnAla	205
Db	3161384	TCCGGGCAAGGAGTCGTGTTCT-----CTCGTCGCCGCC	3161352
Qy	206	SerLysThrIlePheValGlyValAlaAspProAsnAsnPro-----Val	220
Db	3161351	GGCGCACCGTCTACGCGCGTGGGTGACCGGACGCGCACTCGGCGACGGGACCTG	3161292
Qy	221	PheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPhe	240
Db	3161291	TACCGCACGGCGGAC--GGCACGACCTGGGGGCGGTCCCGCGCGCCCTCCCGGACC	3161235
Qy	241	-----IleProHisLysGlyValPheAspProValAsnHisValLeuTyrrIleAla	257
Db	3161234	TCCGCCAAGGTCCCGCTCCGCGCGCGGTACGACACGACACCCCGCGAGCTGTACGTGACG	3161175
Qy	258	ThrSerAsnThrGlyGlyProTyrrAspGlySerSerGlyAspValTrpLysPheSerVal	277
Db	3161174	TACGGCAGCGACCCCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	3161115
Qy	278	ThrSerGlyThrTrpThrArgIleSerProVal-----ProSerThrAspThr	293
Db	3161114	GCCAGGGGACGTGACCGAGGTCACTCCCGGTGAAGCGGGCGGGACGACGACGACGCGC	3161055
Qy	294	AlaAsnAspTyrrPheGlyTyrrSerGlyLeuThrIleAspArgGlnHisProAsnThrIle	313
Db	3161054	TCGGCGACACCTTCGCTACGGCGGGGTCTCGCTCGACGCGCGCGCGCGCGCGCGCGC	3160995
Qy	314	MetValAlaThrGlnIleSerTrpTrpProAspThrIleIlePheArgSerThrAspGly	333
Db	3160994	GTCTGCTCCCAACCAACACCGTGGCGGCGGACACCGGTCTTCGGTCTCCAGGACGGC	3160935
Qy	334	GlyAlaThrTrpThrArgIleTrpAspTrpThrSerTyrrProAsnArgSerLeuArgTyrr	353
Db	3160934	GGCGGTACTGGACGTCCCTCAAGGAC-----GCCGCC	3160902
Qy	354	ValLeuAspIleSerAlaGluProTrpLeuThrPheGlyValGlnProAsnProProVal	373
Db	3160901	GTGTTTCGACGTCTCGAGACTCCCTCTCTCTCGACTGGGGCGACAC-----	3160857
Qy	374	ProSerProLysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAsp	393
Db	3160856	---AAGCCGAACTTCGGCTGGTGGATCCAGGCGCTCGCGCTCGACCGTACGACTCCAG	3160800
Qy	394	ArgMetLeuTyrrGlyThrGlyAlaThrLeuTyrrAlaThrAsnAspLeuThrLysTrpAsp	413
Db	3160799	CACGTCTGTACGGACCGCGCGGACCTCTACGGCACCGGACCTCAAGCCGCTGG---	3160743
Qy	414	SerGlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluThrAlaValAsn	433
Db	3160742	-----GCACCGCGATCCCGCGGCTCGAGGAGAGCGCGCTGCGC	3160704
Qy	434	AspLeuIleSerProProSerGly---AlaProLeuIleSerAlaLeuGlyAspLeuGly	452
Db	3160703	CAACTGATCTCGCCCGGTCGGGAGGACACCTGATCAGCGGACTCGGGGACATCGGT	3160644
Qy	453	GlyPheThrHisAlaAspValThrAlaValProSerThr---IlePheThrSerProVal	471
Db	3160643	GTGATGTACACGACGGGTCTACGGCGTCTCGTTCGCGGCGCATCGGACGACCCCGTG	3160584
Qy	472	PheThrThrGlyThrSerValAlaGluLeuAsnProSerIleIleValArgAla	491
Db	3160583	TTCGGTTCGGCACGGGACTCGCGGAGCGCGCGCGCGCGGTATGTCTCGGACG	3160524
Qy	492	GlySerPheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGly	511
Db	3160523	GGCTGGGCGACACCGGCAACGCG-----GCGTACTCCACGACGCGC	3160482
Qy	512	GlyLysAsnTrp-----PheGlnGlySerGluProGlyGlyValThrThrGlyGlyThr	529
Db	3160481	GGCGGACGTGGCGGCCCTTCGAGGCCACCGCCGACATCGCCAAAGACGACCGCGGACC	3160422
Qy	530	ValAlaAlaSerAlaAspGlySerArgPheValTrpAla-----ProGlyAsp	545
Db	3160421	ATCCGACACGATGCGGCGGCGGACACTGTGTGTCTCTCTGCTGCTGAGGACGCGC	3160365
Qy	546	ProGlyGlnProValValTyrrAla-----ValGlyPheGlyAsnSerTrpAla	561
Db	3160364	-----ACGACGTACCGGGGCCACCGCTCGACGACGACGCGCGGAGCTGGTCC	3160317
Qy	562	AlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThr	581
Db	3160316	GAGGTCTCTCTCTCCGAAAGCGCGCACACCGGTGCGCGACCGCGCGCATCGGACGCGC	3160257
Qy	582	PheTyrrAla-----LeuSerAsnGlyThrPheTyrrArgSerThrAspGlyGlyValThr	599
Db	3160256	TTCTACGCTTACGACTTCGACAAATGACGCTATACGCCGACGACTGACAGTGGCGGTTCG	3160197
Qy	600	PheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHisAla	619
Db	3160196	TTCACGGCGGTGGGCGGAGTGGCC-----TCGGGCGACGCGGAGTTCAGCTGGTGGC	3160140
Qy	620	ValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSer---SerGlyLeuTyrrHisSer	638
Db	3160139	GCGCGGGACGAGGCGGACCTGTGGTCTCCGCCAATGGAACGGGCTCTACCGGTCC	3160080
Qy	639	ThrAsnGlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnValGly	658
Db	3160079	ACCGACGGCGGGACACCTTCGCCAGGATCGACAGCTGTGGGCGCTCTGACACCTCGGC	3160020
Qy	659	PheGlyLysSerAlaProGlySerSerTyrrProAlaValPheValGlyThrIleGly	678


```

Db      3160019  TTCCGGAAGCGCGCCGACCGCGCCGACTACCCCGCGATCTTACCAGGTCCGCTCGACGGAG 3159960
QY      679  GlyValThrGlyAlaTyrArgSerAspCysGlyThrThrTrpValLeuIleAsnAsp 698
Db      3159959  ACCATCAGCGCGCTTACCGCTCCGACGACCGCCGAGACATGGTCCGGATCAACGAC 3159900
QY      699  AspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArg 718
Db      3159899  GACGCCACCAAGTGGGGTGATCGCGAGCGCGTCTCGTGGTGACCCGCGCATCCACGGC 3159840
QY      719  ArgValTyrIleGlyThrAsnGlyArgGlyIleValTyrClyAsp 733
Db      3159839  CCGGTCTACTCTGCCACCAACGGCGCGGCATCCAGTACGCGGAG 3159795

RESULT 9
US-09-927-827-21
; Sequence 21, Application US/09927827
; Publication No. US20030036176A1
; GENERAL INFORMATION:
; APPLICANT: Bower, Stanley G.
; APPLICANT: Ramsesier, Thomas M.
; TITLE OF INVENTION: Directed Genetic Engineering of Xanthomonas campestris
; FILE REFERENCE: 38-10(15824)B
; CURRENT APPLICATION NUMBER: US/09/927,827
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/279,493
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 69
; SEQ ID NO 21
; LENGTH: 3668
; TYPE: DNA
; ORGANISM: Xanthomonas campestris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1001)..(2668)
US-09-927-827-21

Alignment Scores:
Pred. No.:      3,77e-99      Length:      3668
Score:          1139.00      Matches:      267
Percent Similarity: 50.9%      Conservative: 120
Best Local Similarity: 35.1%      Mismatches: 295
Query Match:      28.2%      Indels:      79
DB:               3          Gaps:      21

US-09-917-376-3 (1-740) x US-09-927-827-21 (1-3668)

QY      1  AlaThrGlnProTyrThrTrpSerAsnValAlaIleGlyGlyGlyPheValAsp 20
Db      1217  GCCACGTCCGGGCGCTTACCAGTCCGCGAGTGTCCCATTTGGCGGTGTGTGTCAACC 1276
QY      21  GlyIleValPheAsnGluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGly 40
Db      1277  GGTGTGCTGTTCATCCCGCCGAACTGGTGTCTGCTATGCGCGCACCGCATGTGGGTGGC 1336
QY      41  MetTyrArgTrpAspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrp 60
Db      1337  GCGTACCGTGGATGGCGAGCGGCGACGAGTGGACCGCGCTGACCGACTGTTGGGCGCT 1396
QY      61  AsnAsnTrpGlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLys 80
Db      1397  GACCATGTGAACTGATGGCATCGACGCAATTCGCGTCGACCCCGCATGCCGATGCG 1456
QY      81  ValTrpAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeu 100
Db      1457  CTGTATCTGGCGCCCGGCACTTATATGATGAA---CGCGCGGCGCACTGCCGAGTGTG 1513
QY      101  ArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGly 120
Db      1514  CGCTCGTTCAACCGCGCGCGACGTTTCGACGCTGCGCGACCTGCCGGTTTAACTGGGT 1573
QY      121  AsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAspAsnIle 140

```


Db 655 ACGGGCGGTTCCGGACAGAGCCCGCTCGATCGCGCGGAGCCCATGAAGTCGCT 714
 Qy 247 PheAspProValAsnHisValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyrAsp 266
 Db 715 CTCACCCCC-----AACTTCTCTACGTGACTTACGCCGACTACCTGGTGCATGGCG 768
 Qy 267 GlySerSerGlyAspValTyrPheSerValThrSerGlyThrTrpThrArgIle--- 285
 Db 769 GTCACGTTCCGCGAAGTCTGGCGCAGACCGACCTCGCGCGCTGGGACGACATTACT 828
 Qy 286 -----SerProValProSerThrAspThrAlaAsnAspTyrPheGly 299
 Db 829 CCGCGCTCGCAACTCGTTCCTCCCTGCTACACCAACAGAGCTTCCCTCGCGCGCGA 888
 Qy 300 TyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIle 319
 Db 889 TTTTGGGTCTCAGCGTCGAGCGGACCAACCCCAACCGTCTCGTGTGCATCACC---CTC 945
 Qy 320 SerTrpTrpProAspThrIle-----IlePheArgSerThrAspGlyGlyAlaThr 336
 Db 946 GACCGCGACCCCGACCGCCCTCGACAGCATCTACCTCTCAACCGATCGCGCGGAC 1005
 Qy 337 TrpThrArgIleTrpAspTrpThrSer-----TyrPro 347
 Db 1006 TGAAGACGTCACCCAGCTCTGTCCCTGCTCCACCTCGAAGTAACTGGGGCCACCCG 1065
 Qy 348 AsnArgSerLeuArgTyrValLeuAspIleSerAlaGluProTrpLeuThrPheGlyVal 367
 Db 1066 ACTAACGCGCGGGTAC---AAGGACGCGACCGCTTCTCGTGGCTCGACTTCAACAC 1122
 Qy 368 GlnPro-----AsnProProValProSerPro-----LysLeuGly 379
 Db 1123 GGTCCCGAGTGGGGGGGATACGGTGCAGCGGTACGCGCGCTCCACCAAGTTGGC 1182
 Qy 380 TrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThr 399
 Db 1183 TGTGGATGAGCGTGTGCTTATCGATTATCGATCCCTCAACCCCGACGACCTGTATGAC 1242
 Qy 400 GlyAlaThrLeuTyrAlaThrAsnAspLeuThrLys-----TrpAspSerGly 415
 Db 1243 GGGGCGACCATCTGGGGCGAGCGACACGCTCTCCGTGTCAGAGGACTCG----- 1293
 Qy 416 GlyGlnIleHisIleAlaPro-----MetValLysGlyLeuGluGluThrAla 431
 Db 1294 -----GGCGGAGCTGTGTACTCCAGATCGACGCTATCGAGGAGAATGCG 1338
 Qy 432 ValAsnAspLeuIleSerProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeu 451
 Db 1339 ATCTGTGCTCGCTCGCCCAAGAGCGCGCGCTCTCTGTCGGGCATCGGTGACATT 1398
 Qy 452 GlyGlyPheThrHisAlaAspValThrAlaValProSerThrIlePheThrSerProVal 471
 Db 1399 AGCGCATGAGACAGACGACCTCACCAAG---CCCAAGAGATGTTGTGGCGCCCG 1455
 Qy 472 PheThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAla 491
 Db 1456 TTCTCCAACTCGACACGATCGAGCTGCGGCAACTTCCCAACGTTGTCGTCGCGGCC 1515
 Qy 492 GlySerPheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGly 511
 Db 1516 GGATCTCTCGGGACACGAGTACGACGCGCGTGGCGCGCTGCTGACGCGACTGACGCG 1575
 Qy 512 GlyLysAsnTrp-----PheGlnGlySerGluProGlyGlyValThr----- 526
 Db 1576 GGAGACGCGGGACCATCTTCTCCCTACTCTGCTCCCTGTCATGAACGCGACCTACCCAG 1635
 Qy 527 GlyGlyThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaPro-----Gly 544
 Db 1636 GGCAGACGATGAGTCTGAGCGGAGCGGCGGATCGTGTGTGTCGACCAAGCTTGAC 1695
 Qy 545 AspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGln 564
 Db 1696 GAGCAGGCTCGGGACCGTGTACTCGACGACTATGCGAAGCGTGTCT----- 1746

Qy 565 GlyValProAla-----AsnAlaGlnIleArgSerAspArgValAsn 578
 Db 1747 ---GTTCCCGTGGCGACCTGAAGGCCCCAGACTGCAATGTCTCTCGACAAAGTCCAG 1803
 Qy 579 ProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyVal 598
 Db 1804 GATGGCAGTTCTACGTACCGATGGCGGCAAGTTCTTCGTCGACCGACGCGCGGAG 1863
 Qy 599 ThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHis 618
 Db 1864 TCGTATCCCGCCCAAGGGCGCGGACTTGTCACT-----GGCACATCGCTCATGCTGCC 1917
 Qy 619 AlaValProGlyLysGluGlyAspLeuTrpLeu-----AlaAlaSerSerGlyLeuTyrHis 637
 Db 1918 GTGAACCCCTGGGTGGCGGCGAGCGTCTGGGTGCTGTCGAGGGCGGTCTCTTCCAC 1977
 Qy 638 SerThrAsnGlyGlySerSerTrpSerAlaIle---ThrGlyValSerSerAlaValAsn 656
 Db 1978 TCGACCGACTTGGCGCTCGTTCCAGAGGTAGTAGTACCGCAACGCGACCTCGTGAGC 2037
 Qy 657 ValGlyPheGlyLysSer-----AlaProGlySerSerTyrProAlaValPhe 672
 Db 2038 GTCCGCCCCCAAGTCCAGTCGAGCGGCAAGAGGTACGCGCCCTCCCGGCTCTTC 2097
 Qy 673 ValValGlyThr-----IleGlyGlyValThrGlyAlaTyrArgSerAspCysGly 690
 Db 2098 ATCTGGGGCACCGACAAAGCCCTGGAAGCGACATCGCGCTGTACCGCTCCGACGACACGCG 2157
 Qy 691 ThrThrTrpValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIle 710
 Db 2158 AGCACCTGGAGCGCGCTCAATGACGAGGACCACTACTCGGGC---CCCACCATGATC 2214
 Qy 711 ThrGlyAspHisAlaAsnLeuArgValTyrIleGlyThrAsnGlyArgGlyIleVal 730
 Db 2215 GAGCGCCACCCCAAGGTCTACGGCGCGTGTATCTAGGCAAGCGCGCGGTATCGTG 2274
 Qy 731 TyrGlyAspIle 734
 Db 2275 TACCGCGACCTT 2286
 RESULT 11
 US-10-395-241-11
 ; Sequence 11, Application US/10395241
 ; Publication No. US20040038367A1
 ; GENERAL INFORMATION:
 ; APPLICANT: YAOI, Katsuro
 ; APPLICANT: MITSUISHI, Yasushi
 ; TITLE OF INVENTION: NOVEL XYLOGLYCAN OLIGOSACCHARIDE-DEGRADING ENZYME, POLYNUCLEOTIDE
 ; TITLE OF INVENTION: ENCODING THE SAME, AND METHOD OF PREPARING THE ENZYME
 ; FILE REFERENCE: 073756
 ; CURRENT APPLICATION NUMBER: US/10/395,241
 ; CURRENT FILING DATE: 2003-03-25
 ; PRIOR APPLICATION NUMBER: JP 2002-83433
 ; PRIOR FILING DATE: 2002-03-25
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 11
 ; LENGTH: 2646
 ; TYPE: DNA
 ; ORGANISM: Geotrichum sp. M128
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (120)..(2558)
 US-10-395-241-11
 Alignment Scores:
 Pred. No.: 3,31e-97 Length: 2646
 Score: 1117.00 Matches: 281
 Percent Similarity: 50.4% Conservative: 124
 Best Local Similarity: 35.0% Mismatches: 279
 Query Match: 27.7% Indels: 120
 DB: 7 Gaps: 31

Qy 657 ValGlyPheGlyIysSer-----AlaProGlySerSerTyrProAlaValPhe 672
Db 2226 GTGGGCGCCCAAGTCCAAAGTCGAGCGCAAGAGGCTAGCGCGCTCCGCGGTCTTC 2285
Qy 673 ValValGlyThr-----IleGlyGlyValThrGlyAlaTyrArgSerAspCysGly 690
Db 2286 ATCTGGGGACCCAGCAAGCTTGGAGCGACATCGGCTGTACCGCTCCGACGCAACGGC 2345
Qy 691 ThrThrTrpValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIle 710
Db 2346 AGCACCTGGACGCGCTCAATGACAGGAGACACAACTACTCGGCG-----CCACCATGATC 2402
Qy 711 ThrGlyAspHisAlaAsnLeuArgValTyrIleGlyThrAsnGlyArgGlyIleVal 730
Db 2403 GAGCGCGACCCCAAGGTCCTAGCGCGCGGTGTATCTAGGACGAAACCGCGCGGTATCGTG 2462
Qy 731 TyrGlyAspIle 734
Db 2463 TAGCGCGACCTT 2474
RESULT 12
US-10-395-241-17
; Sequence 17, Application US/10395241
; Publication No. US20040038367A1
; GENERAL INFORMATION:
; APPLICANT: YAOI, Katsuro
; APPLICANT: MITSUISHI, Yasushi
; TITLE OF INVENTION: NOVEL XYLOGLUCAN OLIGOSACCHARIDE-DEGRADING ENZYME, POLYNUCLEOTIDE
; TITLE OF INVENTION: ENCODING THE SAME, AND METHOD OF PREPARING THE ENZYME
; FILE REFERENCE: Q73756
; CURRENT APPLICATION NUMBER: US/10/395,241
; CURRENT FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: JP 2002-83433
; PRIOR FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
; LENGTH: 2481
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Xyloglucan Oligosaccharide-Degrading Enzyme with Histidine Tag
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2481)
US-10-395-241-17
Alignment Scores:
Pred. No.: 5.95e-97 Length: 2481
Score: 1114.00 Matches: 281
Percent Similarity: 50.2% Conservative: 123
Best Local Similarity: 35.0% Mismatches: 280
Query Match: 27.6% Indels: 120
DB: 7 Gaps: 31
US-09-917-376-3 (1-740) x US-10-395-241-17 (1-2481)
Qy 6 TyrThrTrpSerAsnValAlaIleGlyGlyGlyPheValAspGlyIleValPheAsn 25
Db 13 TAGAGTTCAAGAATGTGCGGATCGGCGCGCGGGGTACATTACCGGGATTGTGCGGCAC 72
Qy 26 GluGlyAlaProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrpAsp 45
Db 73 CCAAGACCAAGACCTGTGTCAGCGCGCACGACATTCGGCGCGGTACCGCTGGGAC 132
Qy 46 AlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyr 65
Db 133 GCAGGACGCTCCAGTGGATCCGCTCAACGATTTATCGAGGCGGACGACATGAACATT 192
Qy 66 AsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaVal 85
Db 193 ATGGGCACCGAGTCGATCGGCTGGAGCCCAACCAACCCCGACAGGCTGTACCTCGCGCAG 252

Qy 86 GlyMetTyrThr---AsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAsp 104
Db 253 GGGCGCTATGTGCGCGACAGTGG-----GCGCGCTTCTATGTGTCCGAAGAC 300
Qy 105 GlnGlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGly 124
Db 301 CGCGGCGCATGTTTACAACTTACAGAGTCGCGGTTCCCGATCGGCGCCCAACGACATGGGA 360
Qy 125 ArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGly 144
Db 361 CGCAACATGCGAGCGCTCGTGTCAACCGTTCACTCGAACGAGTCTGGATGGGT 420
Qy 145 AlaProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMet 164
Db 421 ACGCGTACA---GAGGTGTCTGGAAGATTGCGACCGCGCCCAAGACCTGGACAAACGTC 477
Qy 165 ThrAsnPheProAspValGlyThrTyrIleAlaAsnProThrAspThrThrGlyTyrGln 184
Db 478 ACGTCCATCCCGGACGCGTTC-----ACCAACGGTATCGGATACACG 519
Qy 185 SerAspIleGlnGlyValValTrpValAlaPheAspLysSerSerSerSerLeuGlyGln 204
Db 520 TCG-----GTCAATTTTCGACCC-----GAA 540
Qy 205 AlaSerLysThrIlePheValGlyValAlaAspProAsnAsnProValPheTrpSerArg 224
Db 541 CGTAATGCGCACCATCTACGCGAGCGGACTGCCCGCAGGGC---ATGTACGTACGCAC 597
Qy 225 AspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPheIle----- 241
Db 598 GACGCGGTGTCTCGTGGGAGCCAGTGGCGGCGCAGCGCTCCAGCTGGCTCAACAGGACC 657
Qy 242 -----ProHisLysGlyVal 246
Db 658 ACGGCGGTTCGCGGCAAGAACGCCGCTGATCGCGCGCGCAGCCCATGAAGTCGCT 717
Qy 247 PheAspProValAsnHisValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyrAsp 266
Db 718 CTCACCCCC-----AATCTCTCTACGTGATTACGCGGACTACCTGGTTCATGGGCG 771
Qy 267 GlySerSerGlyAspValTrpLysPheSerValThrSerGlyThrTrpThrArgIle--- 285
Db 772 GTCACGTTCGCAAAAGTCTGCGCGCAGAACCGACCTCGGCGCGCTGGGACGACATTACT 831
Qy 286 -----SerProValProSerThrAspThrAlaAsnAspTyrPheGly 299
Db 832 CCGCGCGTCGCAACTCGTCGCTCCCGCTACCAACACGACGCTTCCTCGCGCGGA 891
Qy 300 TyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIle 319
Db 892 TTTTGGCGTCTCAGCGTCAGCGACCAACCCCAACCGTCTCGTCTCATCACC---CTC 948
Qy 320 SerTrpTrpProAspThrIle-----IlePheArgSerThrAspGlyAlaThr 336
Db 949 GACGCGACCCCGGACCGCCCTCGACAGCATCTACCTCTCAACCGATCGCGCGGACC 1008
Qy 337 TrpThrArgIleTrpAspTrpThrSer-----TyrPro 347
Db 1009 TGAAGGACGCTCACCCAGCTCTCGTCCCGCTCAACCTCGAAGGTAACTGGGCGCCACCG 1068
Qy 348 AsnArgSerLeuArgTyrValLeuAspIleSerAlaGluProTrpLeuThrPheGlyVal 367
Db 1069 ACTAACCGCGCGCGGTAC---AAGACGCGACGCTGTGTTCGCTGGCTCGACTTCAACAAAC 1125
Qy 368 GlnPro-----AsnProProValProSerPro-----LysLeuGly 379
Db 1126 GGTCCCGAGTGGGGGATACGCTGCGCGCACGTCAGCGCGGCTCACCAAGTTTGGC 1185
Qy 380 TrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeuTyrGlyThr 399
Db 1186 TGGTGGATGAGCGGTGTGCTTATCGATCCGTTCAACCCCGACGACCTGTATCGCGCAG 1245

Qy	400	GlyAlaThrLeuTyrAlaThrAsnAspLeuThrLys-----TrpAspSerGly	415
Db	1246	GGGGCCACCATTGGGCGCAGCCGACACGCCTCTCCGTGTCGAGAGGACTGG-----	1296
Qy	416	GlyGlnIleHisIleAlaPro-----MetValIysGlyLeuGluThrAla	431
Db	1297	-----CGGCCGAGCTGGTACCTCCAGATCGACCGTATCGAGGAGAATGCG	1341
Qy	432	ValAsnAspLeuIleSerProSerGlyValAlaProLeuIleSerAlaLeuGlyAspLeu	451
Db	1342	ATCCTGTGCTCGCTCGCCCAAGAGCGGCGCGGCGCTCTCTGTGGGCATCGGTGACATT	1401
Qy	452	GlyGlyPheThrHisAlaAspValThrAlaValProSerThrIlePheThrSerProVal	471
Db	1402	AGCGGCATGAAGCACGACGACCTCACCAAG---CCCAGAAGATGTTTGGTGGCCCCAG	1458
Qy	472	PheThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAla	491
Db	1459	TTCTCAACCTCGACAGCATCGACGCTGGCGGGCAACTCCCAACGTTGTGTCGCGGCC	1518
Qy	492	GlySerPheAspProSerGlnProAsnAspArgHisValAlaPheSerThrAspGly	511
Db	1519	GGATCTCGGGACACAGTACGACAGCGGTGGCGCGGTGCTGATGACGACTGACGGC	1578
Qy	512	GlyLysAsnTrp-----PheGlnIysSerGluProGlyGlyValThr-----	526
Db	1579	GGAGCGGTGGACCATCTTCCCTACTGTCGCTCTGGCATGAACGCGACCCTACCAG	1638
Qy	527	GlyGlyThrValAlaAlaSerAlaAspGlySerArgPheValThrAlaPro-----	544
Db	1639	GGCAGACGATTTCAGCTCGACGCGCGGCGAGCAGATCGTGTGTCGACCAAGCTTGAC	1698
Qy	545	AspProGlyGlnProValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGln	564
Db	1699	GAGCAGGCTCGGGACCGTGTGTHACTCGACGACTATGGCAAGCGTGCT-----	1749
Qy	565	GlyValProAla-----AsnAlaGlnIleArgSerAspArgValAsn	578
Db	1750	---GTTCCCGTGGCGACCTGAAGGCCAGACTGCCAATGTGCTCTCGACAGGTCCAG	1806
Qy	579	ProIysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyVal	598
Db	1807	GATGGCAGCTTCTACGCTACCGATGGCGGCAAGTTCTCGTCTCGACCGACGCGGGAAG	1866
Qy	599	ThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHis	618
Db	1867	TCGTATGCGCCCAAGGGCGCGGACTGTGCATCT---GGCAGTCGCTCATGCGCTGCC	1920
Qy	619	AlaValProGlyIysGlyAspLeuTrpLeu---AlaAlaSerSerGlyLeuTyrHis	637
Db	1921	GTGAACCCCTGGGTGGCGCGGACGCTGGGTGCTGTTCCCGAGGGCGGTCTCTCCAC	1980
Qy	638	SerThrAsnGlyGlySerSerTrpSerAlaIle---ThrGlyValSerSerAlaValAsn	656
Db	1981	TCGACCGCATTTGGCGCCCTCGTTTCACGAGGGTAGGTACCGCCCAACGACCCCTCGTAGC	2040
Qy	657	ValGlyPheGlyIysSer-----AlaProGlySerSerTyrProAlaValPhe	672
Db	2041	GTGCGGCCCCCAAGTCCCAAGTCGACGGCAAGAGGTAGCGCGCCCTCCGCGGTCTTC	2100
Qy	673	ValValGlyThr-----IleGlyGlyValThrGlyAlaTyrArgSerAspAspCysGly	690
Db	2101	ATCTGGGGCACCGACAAGCCTTGAAGCGACATCGGCTGTACCGCTCCGACGACACGGC	2160
Qy	691	ThrThrTrpValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIle	710
Db	2161	AGCACTTGACGCGCGTCAATGATCCAGGAGCAACAATACTACTCGGGC---CCCACTGATC	2217
Qy	711	ThrGlyAspHisAlaAsnLeuArgValTyrIleGlyThrAsnGlyArgGlyIleVal	730
Db	2218	GAGCGGACCCCAAGGCTTACGGGCGCGGTATCTTAGGCACGAAACGCGCGGTATCGTG	2277
Qy	731	TyrGlyAspIle	734

Db 2278 TACGCCGACCTT 2289

RESULT 13

```

US-10-420-191-1
; Sequence 1, Application US/10420191
; Publication No. US20040067569A1
; GENERAL INFORMATION:
; APPLICANT: No. US20040067569A1ozymes Biotech, Inc.
; APPLICANT: Rey, Michael W.
; APPLICANT: Zaretsky, Elizabeth J.
; APPLICANT: Haas, Jeffrey A.
; TITLE OF INVENTION: Polypeptides Having Xyloglucanase Activity And Nucleic Acids
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 10210.200-US
; CURRENT APPLICATION NUMBER: US/10/420,191
; CURRENT FILING DATE: 2003-04-18
; PRIOR APPLICATION NUMBER: US 60/373,987
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 5698
; TYPE: DNA
; ORGANISM: Tricoderma reesei
US-10-420-191-1

Alignment Scores:
Pred. No.: 9.49e-74
Score: 878.50
Percent Similarity: 29.9%
Best Local Similarity: 22.1%
Query Match: 21.8%
DB: 7
Gaps: 52
Indels: 793
Mismatchches: 270
Conservative: 119
Matches: 334
Length: 5698

```

QY 93 ----- 93
Db 781 GCGGTACATATGCTGTGCGACACAGACACTACAGAGAGTCTAGATCTCGATACAACTGC 840
QY 94 -----ProAsnAspGlyAlaIleLeuArgSerSerAspGln 105
Db 841 CTGACATTGGCCATATATAGGGATCCGAGTAATGGAGCCATCATTCCTGCTAGACCCG 900
QY 105 ----- 105
Db 901 GACTGTAAACGGTATATATCCCTAGGCTCATTTACCTCGGTAGTAAGCGAGCTGGCG 960
QY 106 GlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyAsnMetProGlyArg 125
Db 961 GCGCAACGCGTCTTCCACCACTTGCCCTTCAAAGTCGGGGTAAACATGCCAGGAGCG 1020
QY 125 ----- 125
Db 1021 CCGGTTGACACAGGAAGTGGTTGAACGGGAAGTTTTCAGCCCCCATTTGACGGTCTCGG 1080
QY 126 GlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAla 145
Db 1081 GGAGCCGGAGAGCTCTGGCTGTGATCGATCCGCGCAACTCCCAACATCATCTACTTTGGTCT 1140
QY 146 Pro----- 146
Db 1141 CTTCCGCCCTCTCGCAGACCGACAGCTAGGCGGTTGAGGTTGTAGTAGATGAACACCA 1200
QY 147 ---SerGlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMet--- 164
Db 1201 CGTCAAGAAACGGCCTCTGGAAGTCTACGGAGCGCGGTGACCTTTTCCAAAGGCTCG 1260
QY 165 -----Thr 165
Db 1261 GCGAGTCCTTTTCGCGGAGACTTTCAGATGCTGCGCGCGCACTCGAAAGGTTCCAGAGC 1320
QY 166 AsnPheProAspValGlyThrTrpIleAlaAsnProThrAspThrThrGlyTyrGln--- 184
Db 1321 TCGTTACGGCAACTGGGAGCTATACATCCAGACCGAGTGATTCCAAACGGCTACACAGC 1380
QY 185 -----Ser 185
Db 1381 AGCAAGTGGCTTGACCTGCATGTAGGCTTGGGCTCCTAAGGTTGCCGATGTTGTGC 1440
QY 186 AspIleGlnGlyValValTrpValAlaPheAspLysSerSerSerLeuGlyGlnAla 205
Db 1441 GACAAGCAAGGACTCATGGGTAGCTTCAGCTCAACACGACAGCAGCCGGGAGCC 1500
QY 205 ----- 205
Db 1501 CTGTTGCTTCTCAGTACACCCCAATCAAGCTGAGTTGGTGGTCTGCTGGCCCTCGG 1560
QY 206 SerLysThrIlePheValGlyValAlaAspProAsnAsnProValPheTrp----- 222
Db 1561 ACGTCTCGTATCTTGTGGACGGCTGATGCTTCACTCAACACGACAGCAGCCGGGAGCC 1620
QY 223 -----SerArg 224
Db 1621 TGCAGAGCATAGAAACACCGTCCGACTATTTAGTGAGCAAGTCAATGATCTCGTGC 1680
QY 225 AspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGlyPheIleProHis--- 243
Db 1681 AATGCCGGTCCACGTGGAGTGTGTACCGGGGAGCCAGGAAATACTTTCTCTACAAG 1740
QY 244 -----LysGlyValPhe 247
Db 1741 TTACGGCCGAGGTGCACCTCAGCATGCGCCCGCTCGTCCCTTTATGAAGAGGTGTT 1800
QY 248 -----AspProValAsnHisValLeuTyrIleAlaThrSerAsnThrGlyGlyPro 264
Db 1801 GCGAACTCGAGCCAGCAGAGAGGCGCTTGTATCTGACCTATTCCGATGGCAGGGCCG 1860
QY 264 ----- 264

Db 1861 CGTTTTCAGCTGCTGCTCTCTCCGGAACATAGACTGGATAAGGCTACCGTGTCCCGGC 1920
QY 265 TyrAspGlySerSerGlyAspValTrpLysPheSerValThrSerGlyThrTrpThrArg 284
Db 1921 TATGATGGCACACTGGCTCAGTGTGGAGGTACGACATTCGAGGGGAACTTTGGAAGAC 1980
QY 285 IleSerProValProSerThrAspThr----- 293
Db 1981 ATACTACCGTGAACCGAGTCACACCTCCATCTGTAACGTCGCCCTTGAACCTTTCTG 2040
QY 294 -----AlaAsnAspTyrPheGlyTyrSerGlyLeuThrIleAsp--- 306
Db 2041 ATCACCCTGTCTCTGATCAGATCTATCTTTGCTTTGGCGCCTTTGGCCTCGATTTG 2100
QY 307 -----Arg 307
Db 2101 TAGTGGGACAGACACCTAGTCTAGATATGAACCGAAACCGCGGACCGGAGCTAAAC 2160
QY 308 GlnHisProAsnThrIleMetValAlaThrGlnIleSerTrpTrpProAspThr----- 325
Db 2161 CAAAGCCAGGAACCTTGTGTTGCTTCTTTGAACCTCTTGGTGGCAGATGCTCAGCTG 2220
QY 326 -----IleIle 327
Db 2221 GTTTTCGGTCTCTGGGAACAACAACGAAGAACTTTGAGAACCCCGGTCTACGAGTCCGAC 2280
QY 328 PheArgSerThrAspGlyGlyAlaThrTrpThrArgIleTrpAspTrpThrSerTrpPro 347
Db 2281 TTTCCGTCGACCGACTCTGGGACAAACATGGAGCCGATCTGGCGTGGCGAGCTATCCG 2340
QY 348 AsnArgSer---LeuArgTyrValLeuAspIleSerAlaGluPro----- 361
Db 2341 AAAGCCAGCTGGCTGAGACC-CTGTTGTACCTCGGGCTAGACCCGACCCGCTCGATAGG 2399
QY 362 -----Trp 362
Db 2400 CACTGAGACCTATTACTACAGCATCTCAGTGAGTCACTCTTAAACGATCCGATGCGAATGG 2459
QY 363 LeuThrPhe-----GlyValGlnProAsn----- 370
Db 2460 CTGACTCTGGATAATGATGCTGATAGCTCAGTCACTAGTGAGAATGCTAGGCTACGTTACC 2519
QY 371 -----ProProValPro----- 374
Db 2520 GACTGACTCTGCTTCAAGACTCCCAAGACCCGCTGGATCAAGAACAACTTTATCGATGTG 2579
QY 374 ----- 374
Db 2580 ATGACTGAGACGAAGTTCTGAGGTTTCTGGGCACCTAGTTCTTGTGTAATAGTACAC 2639
QY 375 -----SerPro-----LysLeuGlyTrpMetAspGluAlaMet 385
Db 2640 TCGAGCGAGTCAACCGTCCGATGGTCTCATCAAGCGCCTCGGCTGGATGATTTAGTCTCTC 2699
QY 386 Ala----- 386
Db 2700 GGCTCGCTCAGTGGCAGGCTACAGAGTAGTTCCGGGAGCCGACCTACTAATCAGAGAG 2759
QY 387 ---IleAspProPheAsnSerAspArgMetLeuTyrGlyThrGlyValAlaThrLeuTyrAla 405
Db 2760 CAGATTGACCCACCGACAGCAACCACTGGCTCTACGGCACCGGATGACAACTCTTTGGC 2819
QY 406 Thr----- 406
Db 2820 GTCTAACTGGTGGTGTGCTGTTGGTGACCGAGATGCGTGGCCTTACTGTTAGAAACCG 2879
QY 407 ---AsnAspLeuThrLysTrpAspSerGlyGlyGlnIleHisIleAlaProMetValLys 425
Db 2880 CGCCAGCATCTCACCACACTGGGACACCGCCCAACATGTGTCAATCCAATCACTGCGCAG 2939
QY 426 Gly----- 426

Db 2940 GCGGTGTAGAGTGGTTCACCTGTGCGCGGTGTGTACACAGTTAGTTAGTACCGTCTG 2999
Qy 427 ---LeuGluGluThrAlaValAlaAsnAspLeuIleSerProSerGlyAlaProLeuIle 445
Db 3000 CGCATCGAGGAATTCTCGTCCAGAACCTGCGCTCTGCACCCGCGGAGGAGTATTG 3059
Qy 446 Ser----- 446
Db 3060 GCGTAGCTCTTAAAGACGAGTCTGTGACCGGAGACGTGGCGCGCTTCCTCGATTAAC 3119
Qy 447 ---AlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThrAlaVal----- 462
Db 3120 CCGCAGTCGAGAGCAACACGGCTTACCTTGGCCAGACAAAGACCTCGGCACATCG 3179
Qy 462 ----- 462
Db 3180 GCGGTGACGCTCTGCTGTTGCCGAAGTGAACCGTGTCTTTGCTGGAGCCCTGTAGC 3239
Qy 463 ProSerThrIlePheThrSerProValPheThrThrGlyThrSerValAspTyrAla--- 481
Db 3240 GCGCAGAGCGTCTGGCAACGCCACATGGGCCACCTCGACGAGCGTCTAGCGCCGG 3299
Qy 482 -----Glu 482
Db 3300 AGCGTCTGCCAGACCCGTTGCGGGTGTACCGGTGGAGCTGCTCGACAGTATGCGGCC 3359
Qy 483 LeuAsnProSerIleValAlaGlySer----- 493
Db 3360 TACTCGGTCAAGAGCGTCTCGCGTGGCAACACCGCGCGGCACGACAGGTGGCCATC 3419
Qy 494 ---PheAspProSerSerGlnProAsnAspArgHisValAlaPhe----- 507
Db 3420 TTGAGCCAGTTCTCGCAGCAGCGCAGCGGTGTGGCGCGGTGTGTCACCGGTAG 3479
Qy 508 SerThrAspGlyGlyAsnTrpPheGlnGlySerGluProGlyGlyValThrGly 527
Db 3480 ACGTCCGACGCGCGCGACGTGGAGCATCGACTACGCGCGGCACACGTCCATGAACGCG 3539
Qy 528 Gly----- 528
Db 3540 GCGAGCTGCGCGCGGTGCACCTCGTAGTGTGTCGCGCGGTGTGAGTACTTGGCG 3599
Qy 529 ---ThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaProGlyAspProGly 547
Db 3600 GCGCAGGTGGCTATTTCGCGCGACGCGACAGATCTCTGTGTCGACCGCTCTGCGGC 3659
Qy 548 -----GlnPro 549
Db 3660 GCGTCCACCGGATAAGCCGCTGCGCTGTGTAGGAGACGAGTGGCGGAGCAGCGCG 3719
Qy 550 ValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsn 569
Db 3720 CTGACAGCTCG---CAGTTCAGGGAGCTTTCCTCGCTCGAGCTCGCGCGCGGC 3776
Qy 570 AlaGln----- 571
Db 3777 GCCGAGTCGCGAGGTCAAGGTCCCGTTCGAAACGAGGAGAGCTCGGACGCGCGCGC 3836
Qy 572 ---IleArgSerAspArgValAsnProLysThrPheTyrAlaLeuSerAsnGlyThr 589
Db 3837 CGGCTCATCGCTCGGACAAAGAACCAACAGCGTCTTCTACGCGCGGTTCGGATCGACC 3896
Qy 590 Phe----- 590
Db 3897 TTTTAGTAGCGAGCTGTTCTTCTGTTGTTCGAGAAAGATGCGGCGGAGCTAGTGTG 3956
Qy 591 ---TyrArgSerThrAspGlyGlyValThrPheGlnProValAlaAlaGlyLeuProSer 609
Db 3957 AAAACGTCACAGGACACCGGACGAGCTTC-----ACGCGCGG---CCCAAG 4004
Qy 610 SerGlyAlaValGlyValMet----- 616
Db 4005 CTGGCAGCGCAGGATGACAGTCTGTTGTTGGCGGTGTCAGTGTGCGCGCGCGGTTTC 4064

Qy 617 -----PheHisAlaValProGlyLysGluGlyAspLeu 627
Db 4065 GACCCGTGCGFCCTCGATCCGGGATATCGCTGTCTCACCCGACCCCGGCGCATGTTG 4124
Qy 628 TrpLeuAlaAlaSerSer----- 633
Db 4125 TATGTCTCGACCGCGGTAGGCCCTATAGCAGAGTGGGTGGTGGCGCCCGTGCAC 4184
Qy 634 -----GlyLeuTyrHisSerThrAsnGlyGlySer----- 644
Db 4185 ATACAGAGCTGCTCGCATATTCCTCCACAGATCGGGCACGACCTTTGGCCAA 4244
Qy 644 ----- 644
Db 4245 GTCTCCACCGCCTGAAGCCGTATAAGCGAGGTGTCTAGCCCGTGTGAAACCGGTT 4304
Qy 645 ---TrpSerAlaIleThrGlyValSerSerAlaValAsnValGlyPheGlyLysSer 662
Db 4305 CAGAGTGGCGGACTCCAACACCTACAGATCGCCCTGGGTGGGCTCAGGCTCGAAC 4364
Qy 663 -----AlaProGlySerSerTyr 668
Db 4365 TGAACCTGTATGCTCTGTTGTGGATGTCTAGCGGAGCCACACCGAGTCCGAGCTTG 4424
Qy 669 ProAlaValPheValGlyThrIleGlyValThrGlyAla-----TyrArgSer 686
Db 4425 ACCTTGGACATACGATCGGCACC-----GGCCGTGAGGGCTCGCTCTACGCCAGT 4478
Qy 687 AspAspCysGlyThrTrpVal----- 694
Db 4479 GGAGACAGCGCGCTCTCTGGAAGCGCTGGCGGCGAGTCCCGAGCGAGATCGCGTCA 4538
Qy 695 -----LeuIleAsnAspAspGlnHisGlnTyrGly----- 704
Db 4539 CCTGTGTCGCGCGGAGGACCTCGACATCCAGGGCTCCCGAGGGCTTCGGCTCCATCGAC 4598
Qy 704 ----- 704
Db 4599 AGCACCAAGTTCGCGGACGCGCTGTAGTCCCGAGGGTCCCGAAGCGAGGTAGCTG 4658
Qy 705 AsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgValTyrIleGlyThr 724
Db 4659 TCGTGTTCAGCGCGCTCGCGCAGCAGCAGCGCGG-----CAAGTCTACGTGGGCACC 4712
Qy 725 AsnGlyArgGlyIleValTyrGlyAspIleGlyAlaPro 738
Db 4713 AACGCGCGGGCGTCTTTTACGCTCAGGCGTCTGTGGCGGCC 4754

RESULT 14

US-10-653-047-7511
; Sequence 7511, Application US/10653047
; Publication No. US20040229367A1
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjarke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 5849.200-US
; CURRENT APPLICATION NUMBER: US/10/653,047
; CURRENT FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US/09/533,559
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/273,623
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7511
; LENGTH: 1103

```

; TYPE: DNA
; ORGANISM: Tricoderma reesei
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1103)
; OTHER INFORMATION: n = A,T,C or G
US-10-653-047-7511

Alignment Scores:
Pred. No.: 1,01e-27 Length: 1103
Score: 392.50 Matches: 114
Percent Similarity: 45.0% Conservative: 36
Best Local Similarity: 34.2% Mismatches: 125
Query Match: 9.7% Indels: 58
DB: 8 Gaps: 10

US-09-917-376-3 (1-740) x US-10-653-047-7511 (1-1103)

Qy 302 GlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTrp 321
Db 3 GGCGCTGGCGCTCGATTGCAAAAGCCAGGAACCTTGTGTGCTTCTTGAACCTCTGG 62
Qy 322 TrpProAspThrIleIlePheArgSerThrAspGlyGlyAlaThrTrpThrArgIleTrp 341
Db 63 TGGCCAGATGCTCAGCTGTTTCGTCACCACTCTGGACACATGGAGCCCGATCTGG 122
Qy 342 AspTrpThrSerTyProAsnArgSerLeuArgTyValLeuAspIleSerAlaGluPro 361
Db 123 GCGTGGCGGAGTATCCGATGAGACCTATTACTACAGCATCTCAACTCCCAAGCACCG 182
Qy 362 TrpLeuThrPheGly---ValGlnProAsnProProValProSer-----Pro 376
Db 183 TGGATCAAGAACAACTTTATTCGATGTGACGAGCGAGTCACCGTCCGATGCTTATCAAG 242
Qy 377 LysLeuGlyTrpMetAspGluAlaMetAlaIleAspProPheAsnSerAspArgMetLeu 396
Db 243 CGCTCGCGTGGATGATGATGCTNTGAGATGACCCACCGGAGGAACTTGGCTT 302
Qy 397 -TyGlyThrGlyAlaThrLeuTyAlaThrAsnAspLeuThrLysTrpAspSer----- 414
Db 303 TTACGGACCGGATGACATNTTGGCGGCCAGCATTTCCCACTTGGACAGCGGCC 362
Qy 415 -----GlyGlyGlnIleHisIleAlaProMetValLys-GlyLeuGluThrAlaV 432
Db 363 ACAATGTGGTCAATCCCAATTACTTGGCAGACGGGATTGGAAGGAATT-----TTCCGT 416
Qy 432 alaAsnAspLeuIleSerProProSerGly----- 441
Db 417 TCAAGGACCTGGCGCTTTTCAACCGGGGGAAGCGAGCTTTTGGCCGCAAGTCCGAGACG 476
Qy 442 -----AlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspV 459
Db 477 ANCAACGGGTTTACCTTTTCCCGCAGCAGAACGACCTTGGG----- 517
Qy 459 alThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThr-SerVal 478
Db 518 -----ACATTGGCCAGACGGTTTGGCAACTCCACATCGGCCACCTCGACGAGGCTC 572
Qy 479 AspTyrla--GluLeuAsnProSerIleIleValArgAlaGlySer----- 493
Db 573 GACTACCGCGGAACTCGTCAAGAGCGGTTTCTCGCGCTCGGCAACACCGCGCGGACG 632
Qy 494 -----PheAspProSerSerGlnProAsnAspArgHisValAlaPheS 508
Db 633 CAACAAGTGGCGCATTTTTCGAAACGGCGGGCGGCGAGCTNGGAA-----GCAATTC 686
Qy 508 erThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGlyValThrThrGlyG 528
Db 687 GAACAACTACGCT-----GGNTCCGAAACACCGCTTCCATTGAACCGCG 728
Qy 528 lyThrValAlaIleSerAlaAspGlySerArgPheValTrpAlaProGlyAspProGlyG 548
Db 729 GCGCGGTGGCTATTTCGGCGCAGCGGCGACACGATCTCTCTGTCGACCGCTCGTCCGGC- 787

```

```

Qy 548 lnProValValTyAlaValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProA 568
Db 788 -----GTGCAGCGCTCGCAGTTTCCAGGGCAGCTTTCCTCTCGAGCTGCCG 839
Qy 568 laAsnAlaGlnIleArgSerAspArgValAsnProLysThrPheTyrlaLeuSerAsnG 588
Db 840 CGGCGCGCTCATCGNCTCGGACAAAGACCAACAGCGTNTTCTACGCGCGCTCCGGAT 899
Qy 588 lyThrPheTyrlaSerThrAspGlyGly 597
Db 900 CGACCTTTTACGTACGACGAGGACACCGCGC 928

RESULT 15
US-10-246-330-3
; Sequence 3, Application US/10246330
; Publication No. US20030166030A1
; GENERAL INFORMATION:
; APPLICANT: O'Toole, George A.
; APPLICANT: Mah, Thien-Fah
; TITLE OF INVENTION: METHODS TO STUDY AND MECHANISMS OF
; FILE REFERENCE: 14537-003001
; CURRENT APPLICATION NUMBER: US/10/246,330
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 60/323,241
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 7407
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(7404)
US-10-246-330-3

Alignment Scores:
Pred. No.: 9,48e-12 Length: 7407
Score: 238.00 Matches: 205
Percent Similarity: 32.8% Conservative: 88
Best Local Similarity: 22.8% Mismatches: 314
Query Match: 5.9% Indels: 292
DB: 6 Gaps: 50

US-09-917-376-3 (1-740) x US-10-246-330-3 (1-7407)

Qy 11 ValAlaIleGlyGlyGlyPheValAspGlyIleValPheAsnGluGlyAlaProGly 30
Db 3250 GTCAACCTGAGCAATGGCAGCAGCTCAGCGC-----ACCGCCGAGCGGCG 3297
Qy 31 IleLeuTyrlaValArgThrAspIleGlyMetTyrlaArgTrpAspAlaAlaAsnGlyArg 50
Db 3298 AGCACGGTGTATCTCACCGAC-----GGCAACGCGCAAT 3330
Qy 51 TrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyrla----- 65
Db 3331 CCGATCGCCGAGGTTCACCGCGCGCGCGCGCGCAACTGACCTACACCGCTCCACGCGG 3390
Qy 66 -----AsnGly---ValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrp 82
Db 3391 ATCCCAACGCTGTGTGTCAAGTGTGTCGCCAGGACGCTCGGTAAACAGCAGCGCG 3450
Qy 83 AlaAlaValGlyMetTyrlaAsnSer-----TrpAspProAsnAsp 96
Db 3451 CCGCGCAGCGGTACCGTTCGATTCAGCGCGCGCGCGCGCGGTGATCAACCGAGCAAC 3510
Qy 97 GlyAlaIleLeuArg---SerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuPro 115
Db 3511 GCGCTCGTTCATCAGCGGCGCGCGCGCGCGCGCGGTGCGACGCTCACC----- 3561
Qy 116 PheLysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGlyLeuAlaValAspPro 135

```


Db 3562 ---GATCGCGGCAACCCGATAGCGGAGGTCAACCGCGCAC--- 3600
QY 136 AsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLysGlyLeuTyrArgSerThr 155
Db 3601 ---GGCAGCGGCAACTGGAGTTTCAG 3624
QY 156 AspSerGlyAlaThrTyrSerGlnMetThrAsnPheProAspValGlyThrTyrIleAla 175
Db 3625 CGG---GGCAGCCCGCTGGCC 3642
QY 176 Asn---ProThrAspThrThrGlyTyrGlnSer 185
Db 3643 AACGCACGGTGATCGTCGCACCGCCACCGACCGCGCAATACCGCCCGCGAGGCC 3702
QY 186 ---AspIleGlnGlyValValTyrValAlaPheAspLysSerSerSer 201
Db 3703 GCCACACGGTGGAGCGGTGGCGCGCGCGCGGTGATCGATCCGAGCAACGGCAGC 3762
QY 202 Leu---GlyGlnAlaSerLysThrIlePheValGlyValAlaAspProAsn---Asn 218
Db 3763 ACCATCAGCGCACCGCGGAGCGCGGCGCAAGGTGATCTCACCGCAGCGCAACGGCAAC 3822
QY 219 ProVal---PheTyrSerArgAspGlyGlyAlaThrTyrGlnAlaValProGlyAlaPro 237
Db 3823 CGGATCGGCGGAACACCGCGGACGGCAGCGCGCAACTGGAGTTTCACCGCCCGCGCAGCGC 3882
QY 238 ---ThrGly--- 239
Db 3883 CTGGCCAACGCGACGGTGGTCAACGCGCGTGGCCAGGACCGTGGCGCAATACCGGCCCG 3942
QY 240 ---PheIleProHisLysGlyValPheAspProValAsn 251
Db 3943 CAGGCGACGACTACCGTGGACGCGGTGGCGCGCAACCGCTGTGGTCAATCCGAGCAAC 4002
QY 252 HisValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyr---AspGlySer 268
Db 4003 GCGAACCTGCTCAACGGTACCGCGGAGCGCGGCGCAGCAGCGTGCACCTTGACCGCAGCGCAAC 4062
QY 269 SerGlyAspValTyrLysPheSerValThr---SerGlyThrTrp---ThrArgile 285
Db 4063 GCGAACCGGATCGGCGCAGCACCGCCGATGCGACGCGCAACTGGAGTTTCACCGCCCGC 4122
QY 286 SerProValProSer---ThrAspThrAlaAsnAspTyrPheGlyTyrSer 301
Db 4123 TCGCACTGCCCAACGCGCACCGTGTCACTGACGTGACCGGAGCGCGCGCGCGCAATACC 4182
QY 302 GlyLeu---ThrIleAspArgGlnHisProAsnThrIleMetVal--- 315
Db 4183 AGCCTTCCCGCTACACGACGGTGGATTCCTCGCTGCGTGCATCCCGCAGGTGGATCCG 4242
QY 316 ---AlaThrGlnIleSerTyrTrpProAsp---ThrIleIlePheArgSer 330
Db 4243 AGCAACGGTTCGGTGATCAGCGGCGCACCGCGGACCGCGGCAACATCATCATC--- 4296
QY 331 ThrAspGlyGly---AlaThrTrpThrArgIleTyrAsp 342
Db 4297 ACCGATGGCAACGCAACCCGATGGCCAGGTCAACCGCGGACGGCGGCAACTGGTCC 4356
QY 343 TrpThr---SerTyrProAsnArgSerLeuArgTyrVal 354
Db 4357 TTCCTCCAGGCATCCCGCTCCCGATGGCGATGGCGCGGTGTCACGTGGTGGCGCGCAGCCCA 4416
QY 355 LeuAspIleSerAlaGluProTyrLeu---ThrPheGlyValGlnProAsnPro 372
Db 4417 AGCAATCTCAGACGTGCGCGCGGTGATCACTGTGTGATGGTGGTGGCGCGCGCGCGC 4476
QY 373 ValProSerProLysLeuGlyTyrMetAspGluAlaMetAlaIleAspProPheAsnSer 392
Db 4477 GTG---ATCGATCCGAGCAACGCGC 4497
QY 393 AspArgMetLeuTyrGlyThr---GlyAlaThrLeuTyrAlaThrAsnAspLeu 409

Db 4498 ACCGAG---ATAAGCGGTACCGGAGGCGCGCGGCGGATCTCTCACC--- 4545
QY 410 ThrLysTyrAspSerGlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluGlu 429
Db 4546 ---GATGGCGCGCAAC---CCGATCGCGGAGCGCACCGCGCAC 4584
QY 430 ThrAlaValAsnAspLeuIleSerPro---ProSerGlyAlaProLeuIle 445
Db 4585 GGCAGCGGCAACTGGAGTTTCACCGCGGACACCCCGCTGCCCAACCGCGATCAAC 4644
QY 446 SerAlaLeuGlyAspLeuGlyGlyPheThrHis---AlaAspVal 459
Db 4645 GCGGTGGCCAGGACCGCGCGGCAATACAGCGGTTCGCGCAGCGTCACCGTCGATGCC 4704
QY 460 ThrAlaValProSerThrIlePhe---ThrSerProValPheThrThrGlyThrSer 477
Db 4705 ATCGCCCGCGCGCGGTGATCAATCCGAGCAATGGAGTCTGTCATCAGCGGTACG--- 4761
QY 478 ValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySerPheAspProSer 497
Db 4762 ---GCGGAAGCGCGCGCGGTGATCTCACCGACGCAACCGCAACCGCGATC 4812
QY 498 SerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTyrPhe--- 516
Db 4813 GGCCAG---GTCACGCGCAGCGGACGCGCAACTGGAGCTTC 4851
QY 516 --- 516
Db 4852 AGCCCGCGCACCGCGTGGCCAAACGGCTCGGTGATCAATGCGCTGGCCAGGACGCGCC 4911
QY 517 ---GlnGlySerGluProGlyGlyValThr--- 525
Db 4912 GGCAACAAACAGCAGTCCCAACAGCGCACCGCTGCTGCGCGCAGCAGCCCGGTG 4971
QY 526 ---ThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheValTyr 541
Db 4972 ATCGATCCGAGCAACCGGTAGCTGCGGTGATCGCGGTACCGCGGAGCTGGTGCACGCGATC 5031
QY 542 AlaProGlyAspProGlyGlnProVal---ValTyrAlaValGlyPheGlyAsnSer 559
Db 5032 CTCACCGAGCGCAACCGGCAACCGGATCGCGCAGGTCAACCGCGATGGCAGCGCAAC--- 5088
QY 560 TrpAlaLeuThrGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnPro 579
Db 5089 TGGAGCTTCACGCCCGCGCAGCGCG--- 5112
QY 580 LysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArg--- 592
Db 5113 ---CTGTCCAATGGCAGCGTGGTCAATGCGGTGGCGCCAGGACGCTGCC 5157
QY 593 ---SerThrAspGlyGlyVal---ThrPheGlnProValAlaAlaGlyLeuPro--- 608
Db 5158 GGCAACACGAGCGCGCGGTGATCGACCAACCGGTGGCGCGGTGGCGCGCGCGCGGTG 5217
QY 609 ---SerSerGlyAlaValGlyValMetPheHisAlaValProGlyLysGlyGlyAsp 626
Db 5218 ATCGACCGCGAGCAACCGGTGTCGAATCAGCGCACCGCGCAACCGCGGTTCGGGTGATC 5277
QY 627 LeuTyrLeuAlaAlaSerSerGlyLeuTyrHisSer---ThrAsnGlyGlySerSerTrp 645
Db 5278 CTCACCGATGGCAATGGCAATCCGATCGCGCAGACACCTTTCGCGCAGCGCGCAACTGG 5337
QY 646 SerAlaIleThrGlyVal---SerSerAlaValAsnValGlyPheGlyLys 661
Db 5338 AGCTTCACCGCGCGCGCGGTGGCGCAACCGCGCGGTGGTCAATGCGGTGGCGCGCGGAC 5397
QY 662 SerAlaProGlySerSerTyrProAlaValPheValGlyThrIleGlyGlyValThr 681
Db 5398 CCGCGCGGCAATACCG 5457
QY 682 GlyAlaTyrArgSerAspAspCysGlyThrThrTyrValLeuIleAsnAspAspGlnHis 701
Db 5458 ---CCGCTGATCAATCCGACCAAC--- 5478

Qy	702	GlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgArgValTyr	721
Db	5479	-----GGCAGCGTGATCACCGGCACC---GCCGAGGTCGGCGCCCAAGGTG	5520
Qy	722	IleGlyThrAsnGlyArgGlyIleValTyrGlyAspIleGlyGlyAlaProSerGly	740
Db	5521	ATCTCACCGATGGCAACGGCAACCCGATCGCGGAGACCAACCGCCGACGGCAGTGGT	5577

Search completed: March 2, 2006, 20:08:50
 Job time : 11752.8 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 2, 2006, 12:24:59 ; Search time 794.098 Seconds
(without alignments)
2043.020 Million cell updates/sec

Title: US-09-917-376-3

Perfect score: 4036

Sequence: 1 ATTQPTWSNVAIGGGFVD.....YIGTNGRIGIVDGGAPSG 740

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7218512 seqs, 1096188070 residues

Total number of hits satisfying chosen parameters: 14437024

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/abss/ABSSWEB_spool/US09917376/runat_02032006_091502_8293/app_query.fasta.1
-DB=Published_Applications_NA_New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abs03p
-USER=US09917376 @CGN_1_1067 @runat_02032006_091502_8293 -NCPU=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_NA_New:

1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	192	4.8	9903	12 US-11-052-554A-517	Sequence 517, App
2	182	4.5	11151	12 US-11-052-554A-525	Sequence 525, App
3	180	4.5	7231	12 US-11-136-527-2622	Sequence 2622, App
4	178	4.4	9474	12 US-11-052-554A-526	Sequence 526, App

5	175.5	4.3	7572	12 US-11-052-554A-517	Sequence 517, App
6	168.5	4.2	15876	12 US-11-052-554A-660	Sequence 660, App
7	166.5	4.1	2943	12 US-11-052-554A-401	Sequence 401, App
8	165	4.1	6615	12 US-11-052-554A-518	Sequence 518, App
c 9	161	4.0	37507	8 US-10-522-037-2	Sequence 2, Appli
c 10	157	3.9	3155	8 US-10-793-626-3881	Sequence 3881, Ap
c 11	155.5	3.9	2733	8 US-10-793-626-3107	Sequence 3107, Ap
c 12	155	3.8	2595	12 US-11-136-527-3719	Sequence 3719, Ap
c 13	153	3.8	3050	8 US-10-793-626-3661	Sequence 3661, Ap
c 14	152	3.8	3748	8 US-10-793-626-3970	Sequence 3970, Ap
c 15	151.5	3.8	3600	8 US-10-793-626-4165	Sequence 4165, Ap
c 16	150	3.7	3921	12 US-11-052-554A-523	Sequence 523, App
c 17	149	3.7	4101	8 US-10-821-234-579	Sequence 579, App
c 18	149	3.7	5086	12 US-11-186-284-30	Sequence 30, Appl
c 19	149	3.7	5086	12 US-11-091-939-329	Sequence 109, App
c 20	147.5	3.7	2268	8 US-10-517-939-329	Sequence 329, App
c 21	147.5	3.7	2937	9 US-11-202-566-8	Sequence 8, Appli
c 22	147	3.6	4551	8 US-10-220-824-7	Sequence 7, Appli
c 23	146	3.6	2850	12 US-11-052-554A-390	Sequence 390, App
c 24	146	3.6	104299	12 US-11-000-688-1364	Sequence 1364, Ap
c 25	145	3.6	2145	8 US-10-793-626-569	Sequence 569, App
c 26	145	3.6	2283	12 US-11-052-554A-452	Sequence 452, App
c 27	144.5	3.6	191684	12 US-11-121-086-2	Sequence 2, Appli
c 28	143.5	3.6	5453	12 US-11-091-883-108	Sequence 108, App
c 29	143.5	3.6	153376	12 US-11-121-086-5	Sequence 5, Appli
c 30	143.5	3.6	172543	12 US-11-121-086-6	Sequence 6, Appli
c 31	141.5	3.5	3024	8 US-10-517-939-83	Sequence 83, Appli
c 32	141	3.5	2425	8 US-10-793-626-4404	Sequence 4404, Ap
c 33	140.5	3.5	2868	12 US-11-052-554A-562	Sequence 562, App
c 34	140.5	3.5	4386	12 US-11-052-554A-662	Sequence 662, App
c 35	140.5	3.5	88421	12 US-11-205-109-1	Sequence 1, Appli
c 36	140.5	3.5	120096	12 US-11-121-086-24	Sequence 24, Appli
c 37	139.5	3.5	4767	12 US-11-052-554A-659	Sequence 659, App
c 38	139	3.4	11447	12 US-11-186-284-25	Sequence 25, Appli
c 39	139	3.4	11554	12 US-11-169-041-34	Sequence 34, Appli
c 40	138.5	3.4	2179	6 US-09-925-065A-683197	Sequence 683197,
c 41	138.5	3.4	5706	12 US-11-052-554A-519	Sequence 519, App
c 42	138.5	3.4	168516	12 US-11-121-086-3	Sequence 3, Appli
c 43	137.5	3.4	4311	12 US-11-052-554A-524	Sequence 524, App
c 44	137	3.4	2155	12 US-11-122-329-122	Sequence 122, App
c 45	137	3.4	4605	12 US-11-102-476-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-11-052-554A-517
; Sequence 517, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: Patentn version 3.3
; SEQ ID NO 517
; LENGTH: 9903
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-517

Alignment Scores:
Pred. No.: 0.000114
Score: 192.00
Length: 9903
Matches: 194
Percent Similarity: 35.0%
Conservative: 94
Best Local Similarity: 23.6%
Mismatches: 323

Query Match:	4.8%	Indels:	213
DB:	12	Gaps:	43
US-09-917-376-3 (1-740) x US-11-052-554A-517 (1-9903)			
QY	10	AsnValAlaIleGlyGlyGlyPheValAspGlyIleValPheAsnGluGly	27
DB	7399	AACCGCAACATCGGCGCGCAAAACATCGGCAC	7449
QY	28	-----AlaProGlyIleLeuTyrValArgThrAspIleGly	41
DB	7450	AACACCGGTTCGGGGCTTAACCGCGGTCTCAACACATCGGTATCGGCAACACCGGCAAC	7509
QY	42	TyrArgTrpAspAlaAlaAsn---GlyArgTrpIleProLeuLeuAspTrpValGlyTrp	60
DB	7510	TACAACATCGGTTCGGCAACACCGGTAACATAACAAC	7554
QY	61	AsnAsnTrpGlyTyrAsnGlyVal---ValSerIleAlaAlaAsp	74
DB	7555	GGCAACACCGGCAACACACATCGGCATCGGCCTGTCCGGCGCAACACGATCGGGTTC	7614
QY	75	---ProIleAsnThrAsnLysValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAsp	93
DB	7615	GGCCCGGTGAACCGCGC-----ATGCCCAACATGGGCCCTGTTC---AACCTGGCGCAC	7665
QY	94	ProAsnAspGlyAlaIleLeuArgSerSerAspGlnGlyAlaThrTrpGlnIleThrPro	113
DB	7666	AACAACATTGGC-----	7677
QY	114	LeuProPheLysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaVal	133
DB	7678	-----ATGCCCAACCGGCGCAACTTCAACACGAGGCAATTGCC	7713
QY	134	AspProAsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGlyLys---GlyLeuTrp	152
DB	7714	AACCGCGCAACACACATCGGCTTGTTCACACCGGCAACACACGTCGGCATCTCG	7773
QY	153	ArgSerThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThr	172
DB	7774	CTGACCGCGCAGCGGTTCGGGCTTCAGTCCCTGAACTCCGGCGCGGCAACACCGGT	7833
QY	173	TyrIleAlaAsnProThrAspThrThrGlyTyrGlnSerAspIleGlnGlyValValTrp	192
DB	7834	TTCTTCAACTCCGGCACCGGCAACACCGGC-----	7863
QY	193	ValAlaPheAspLysSerSerSerSerLeuGlyGln---AlaSerLysThrIlePheVal	211
DB	7864	---TTGTCAACTCCGGCACCGGCAACACCGGCTTGTTCAACTCCGGCACCGGCAACGTC	7920
QY	212	GlyValAlaAlaProAsnAsnProValPheTrpSerArgAspGlyGlyAlaThrTrpGln	231
DB	7921	GGCATCGGCAACATCGGCACCGGCGGTTCGGCTCGGCCTATCCGGCGACGCGGCGT	7980
QY	232	AlaValProGlyAlaProThrGlyPheIleProHisLysGlyValPheAspProValAsn	251
DB	7981	GGCATCGGCGGCAACCAACTCGGCAGTTTC---AACATCGGCTTGTTCATCGGCGCAC	8037
QY	252	HisValLeuTyrIleAlaThrSerAsnThrGlyGlyProTyrAspGlySerSerGly-As	271
DB	8038	GGCAATGTCGGCATCGGCAACTCGGCGCACCGGCAACGTCGCATCGGCAACACCGGCGCAC	8097
QY	271	pValTrpIlePheSerValThrSerGlyThr-----	282
DB	8098	GGCAACACCGGCATCGGAACAGCGGCAACTTCAACACCGGCTTGTTCACCGGCGGT	8157
QY	282	p-----ThrArgIleSerProValProSerThrAspThrAlaAsnAs	296
DB	8158	GTCAACACCGGCATCGGCAACCGGCGCAACCAACACCGGCTTGTTCACATCGGCGCAC	8217
QY	296	pTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAl	316
DB	8218	TTCAACACCGGCATCGGCAACCGGCGCACTTCAACACCGGCTCTTCAACACCGCGGTAGC	8277
QY	316	aThrGlnIleSerTrp---TrpProAspThrIleIle---PheArgSerThr	331
DB	8278	TACAACACCGGCATCGGCAACCGCGGACATCGGCGCGGTTCATCATCGGCGCAC	8337
QY	332	-----AspGlyGlyAlaThrTrpThrArgIleTrpAspTrpThrSer---	345
DB	8338	ATGAACAAACGGCTTGTCTGTGGCGCGCCGACCGGCGGCTGTGGCGGCAACATACACC	8397
QY	346	-----TyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaG1	360
DB	8398	ATCACCATCGAGCGACCTGCGCGGTCTCTCATGTGCGCATCCCGGTCAACATCCCATC	8457
QY	360	uProTrpLeuThrPheGlyValGlnProAsnProValProSerProLysLeuGlyTr	380
DB	8458	ACCG---CGGACATCAACCAATGTCTCCATCCCGGCATTCAGTTCGCCAGA-ATC----	8508
QY	380	pMetAspLualaMetAlaIleAsp---ProPheAsnSerAspArgMetLeuTyrGlyTh	399
DB	8509	---GACCCAGCGGAGACGTCGACATAGGCATCTCTAGTGGCACCGTCTTGGCCCCGGT	8564
QY	399	rGly---AlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlnI1	418
DB	8565	CGTCCGATCACTTCATGCGGCGGACGCGTCGCGCGCGCGTGGACACACCCATCGAAAT	8624
QY	418	eHisIleAlaProMetValLysGlyLeuGluThrAlaValAsnAspLeuIleSerPr	438
DB	8625	TGACTTCGGCCCC-----TCGCGGCGGATCACTCAACATCGGCAA	8666
QY	438	oProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAs	458
DB	8667	GCCGACGCGTCCACCGTGATCAACATGTCGGCGCGCGCGC-----	8709
QY	458	pValThrAlaValProSerThrIlePheThrSerProValPheThrGlyThrSerVa	478
DB	8710	-----GCCGCGCGGATC-----AGCAT	8726
QY	478	lAspTyrAlaGluLeuAsnProSerIleValArgAlaGlySerPheAspProSerSe	498
DB	8727	TCCGATCATCACTTCGGCGCGCGC-----CCCGGCTTCTCAACGCGCACAC	8774
QY	498	rGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnI	518
DB	8775	CGGCGCG-----TCGTGGCGGCTTCTCAACTCGGCGGTCTCG	8810
QY	518	ySerGluProGlyGlyValThrThrGlyGlyThr-----ValaAlaI	532
DB	8811	CAGCGCATCGGCTTGTCAAACTTCGCAACAACTCGGCGCTCTCAACTCGCCACTAG	8870
QY	532	aSerAlaAspGlySerArgPheVal-----	544
DB	8871	CAGCATGGGAATTCGGGCTTCCAAACTATGGGTGCGTGCAGTCGGGTGGCG-----	8925
QY	544	yAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSer-----Tr	560
DB	8926	-----AATTGGGCAACAGCATCTCGGCGCATCTA	8954
QY	560	pAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLy	580
DB	8955	CAACACCGGCTTGGGAGCACCGGCAAAATGTC-----TCGGGCTTCTCAACATCGG	9005
QY	580	sThrPheTyrAla-----LeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyVa	598
DB	9006	CACCAACTCGGCTGGGTGGTTCAGAACCGGCGC-----	9039
QY	598	lThrPheGlnProValAlaAlaGlyLeu-ProsSerGlyAlaValGlyValMetPhe-	617
DB	9040	-ACGAGACGACCTTCAGCTGGGCTGGCCCAACTCGGCTCT---GGAACTCGGTAG	9095
QY	618	-----HisAlaValProGlyLysGluGlyAspLeuTrpLeuA	630
DB	9096	CGCAACATCGGCAACTACAACCTGGGCGGCGCAACATCGGCGTCTCAACCTGGGCGAG	9155
QY	630	la-----AlaSerSerGlyLeuTyrHisSerThrAsnGlyGlySerSerTrpSerA	647

```

Db 9156 CGCCAAACATCGGCGACTTCAACCTGGCGAGCCGCAACATCGGCGACTTCAACCTGGCGAG 9215
Qy 647 lailethrGlyValSerSerAlaValAsnValGlyPheGlyLysSerAlaProGlySers 667
Db 9216 CGCCAAACA-----TCGGCAGCTCC-AACATCGGCTTCGGCAACGTCGGTCCGGGG--- 9264
Qy 667 eTyrProAlaValPheValValGlyThrileGly-----GlyValThrGlyAlaTyrArgs 686
Db 9265 -----CTCAGCGCGCGCCATCGGCAACATCGGCTTCGGCAACACCGGAAACGGAACA 9316
Qy 686 eAspAspCysGlyThrThrTrpValLeuileAsnAspAspGlnHisGlnTyrGlyAsnT 706
Db 9317 TCGGCATCGGCATACCGCGCACCGGCAACATCGGCTTCGGCAACACCGGAAACGGAACA 9376
Qy 706 rpGlyGlnAlaileThrGlyAsp--HisAlaAsnLeuArgValTyrile-GlyThrA 725
Db 9377 TCGGCATCGGCTGACCGCGGCACACCATGACCGGCTTCGGCGGTGGAACATCGGCGCACCG 9436
Qy 725 snGlyArgGlyileVal-----TyrGlyAspIleGlyGlyAlaProSerGly 740
Db 9437 GCAACATCGGCGTATTCAACTCCGGCACCGCAACATCGGCTTCGGCAACATCGGCGC 9492

RESULT 2
US-11-052-554A-525
; Sequence 525, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 525
; LENGTH: 11151
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-525

Alignment Scores:
Pred. No.: 0.000763 Length: 11151
Score: 182.00 Matches: 187
Percent Similarity: 32.6% Conservative: 83
Best Local Similarity: 22.6% Mismatches: 325
Query Match: 4.5% Indels: 236
DB: 12 Gaps: 41

US-09-917-376-3 (1-740) x US-11-052-554A-525 (1-11151)
Qy 6 TyrThrTrpSerAsnValAlaileGlyGlyPheValAspGlyileValPheAsn 25
Db 4190 TGGCGCTTGGCAATCTCGGACGCGGAAACGTCGGGTTCGGCAACATCGGTCCGGCAACG 4249
Qy 26 GluGlyAlaProGlyileLeuTyrValArgThrAspIleGlyGlyMetTyrArgTipAsp 45
Db 4250 TCGGGTTCGCA-----ACTCGGTCGCG----- 4273
Qy 46 AlaAlaAsnGlyArgTipIleProLeuLeuAspTrpValGlyTyrAsnAsnTrpGlyTyr 65
Db 4274 -----CGGTGGGCC-----TGGCGCGCTTCGGCAACGTCGGGT--- 4306
Qy 66 AsnGlyValValSerileAlaAspProIleAsnThrAsnLysValTrpAlaAlaVal 85
Db 4307 -----TGAGCAATCGCGGACGCAACACTGGGGCTGGCCAAACCTGG 4348
Qy 86 GlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaileLeuArgSerSerAspGln 105

```

```

Db 4349 GTGTGGGCAACATCGGTTGG---CAAACACCGGCA-----CGGGCAACATCG 4393
Qy 106 GlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArg 125
Db 4394 GGATCGGGCTGTTCGGCGACTACCAAGCCGATCGGC-GGCCTCAAC-----TCG 4443
Qy 126 GlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAla 145
Db 4444 GGTAGTGGCAATATCGGATTGTTCATTCCGGCACCGGCAATGTCGGGTTCCTTCAACACC 4503
Qy 146 ProSerGlyLys---GlyLeuTrpArgSer-----ThrAspSer 157
Db 4504 GGCACCGGCAACTTCGGACTGTTCAACTCCGGTAGTTTCAACACCGGCAATCGGTAATAGC 4563
Qy 158 GlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyrileAlaAsnPro 177
Db 4564 GGAACCGGCACTACGGGCTCTTCATGCGGCAATTTCAACACCGGATCGCAACCCC 4623
Qy 178 ThrAspThrThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAspLys 197
Db 4624 GGG-----TCGTACAACACG 4638
Qy 198 SerSerSerSerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspProAsn 217
Db 4639 GGCAGCTTCAATGTCGGT-----GATACCAAC 4665
Qy 218 AsnProValPheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaPro 237
Db 4666 ACCGGTGGTTTCAACCCGGGCGACATCAACCCGGCTG-----TTCAAC 4710
Qy 238 ThrGlyPheIleProHisLysGlyValPheAspProValAsnHisValLeuTyrileAla 257
Db 4711 ACCGGC-----ATTATGAATACGGC 4731
Qy 258 ThrSerAsnThrGlyGlyProTyrAspGly-----SerSerGlyAspValTrpLysPhe 275
Db 4732 ACCGCAACACCGGCGCTCATGTCCGGGACCGACAGCACCGCATGCTGGCGC--- 4788
Qy 276 SerValThrSerGlyThrTrpThrArgIleSerProValProSerThrAspThrAlaAsn 295
Db 4789 -----GGC 4791
Qy 296 AspTyr-----PheGlyTyrSer---GlyLeuThrIleAspArgGlnHisProAsn 311
Db 4792 GACCAGAGGGCGCTTTCGGCTGTCTATGTCATCAGATC----- 4833
Qy 312 ThrIleMetValAlaThrGlnIleSerTrpTrpProAspThrIleIlePheArgSerThr 331
Db 4834 -----CCGCAATTCCCGATCCGATCCGATCCACCG 4860
Qy 332 AspGlyGlyAlaThrTrpThrArgIleTrpAspTrpThrSerTyrProAsnArgSerLeu 351
Db 4861 ACTGGCGGTATCGGCGCCCATCGTCATCCGGACACACGATCCTTCCGCGCTGCACCTG 4920
Qy 352 ArgTyrValLeuAspIleSerAlaGluProTrpLeuThrPheGlyValGlnProAsnPro 371
Db 4921 CAG-----ATCACCAGGCGACGCGGACTACAGCTTCACCGTG---CCGACATC 4965
Qy 372 ProValProSerProLysLeuGly-----TrpMetAsp 382
Db 4966 CCCATCCCGGCATCCACATCGGATCAATGGCGTCGTCCGCTTCACCGGCTTCACCGCCCCG 5025
Qy 383 GluAlaMetAlaileAspProPheAsnSerAspArgMetLeuTyrGlyThrGlyAlaThr 402
Db 5026 GAAGCCACCTGCTGTCCGCTCGGCTGAAGAATAACGGTAGTTCATCAGCTTCGGCCCCATC 5085
Qy 403 LeuTyrAlaThrAsnAspLeuThrLysTrpAsp----- 413
Db 5086 ACGCTCTCGAATATGATATTCGCCGCCATGGATTTCAGTTAGGCGCTGCCCGTCTTGGT 5145
Qy 414 -----SerGly-----GlyGlnIleHisIleAlaProMetValLysGlyLeuGlu 429

```

Db 5146 CCTATACGGCGCAACTCGGACCAATTCATCTTGAGCCATCGTGGT-GGCCGGGATCGG 5204
Qy 430 ThrAlaValAsnAspLeuIleSerProSerSerGlyAlaProLeuIle----- 445
Db 5205 TGTGCCCTCGAGATCGAGCCATCCGCT-...GGATCGCATTCGTTGAGTGAGTCGAT 5261
Qy 446 -----SerAlaLeuGlyAspLeuGlyGlyPheThrHisAla 457
Db 5262 TCCTATCCGCATACCTGTTGATATTCGGCTCGGT-...ATCGATGGATTTCAATGTCG 5319
Qy 458 AspValThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSer 477
Db 5320 GAAGTGTGCGATCGATGTCGTCGTCGACATCCCGCGGTTCAGATCAGACGACCCACC 5379
Qy 478 ValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySer----- 493
Db 5380 ATTTCCCGCATCCGCTGGCTTCGACATTCGCACCATGTCGCGACCCCTCAACATCCCG 5439
Qy 494 ---PheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGly 512
Db 5440 ATCATCGACATCCGCGCGCGCGCGG-...TTCCGGGAACTCGACCCAGATCCG 5490
Qy 513 LysAsnTrpPheGlnGlySerGluProGlyValThrThrGlyValThrValAlaAla 532
Db 5491 TCGTCGGGTTCTCAACACGGTCCGCGCGCGGATCGGCGATCGGCAACTCGGTCG 5550
Qy 533 SerAlaAspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProValValTyr 552
Db 5551 GGCCTGTCGGG-...CTGCTCAACAGCGCGCGCGGTCTACCTGTTGGGAC 5601
Qy 553 AlaValGlyPheGlyAsnSerTrpAlaAspGlnGlyValProAlaAsnAlaGlnIle 572
Db 5602 CTCTCGGGGTCGGAATCGCGCACCTCGCTCGGCTGCTGAACTCCGCGCACCGCC 5661
Qy 573 ArgSer-----AspArgValAsnProLysThrPheTyrAla 584
Db 5662 ATCTCCGGGCTGTTCAACGTGAGCACCTCGAGCGCCACCCCGCGGTGATCTCGGG 5721
Qy 585 LeuSerAsn-----GlyThrPheTyrArgSerThrAspGlyGlyVal 598
Db 5722 TTCAGCAACCTCGGCGACCATATGTCGGGGTG-----TCCATCGATGGCTGATC 5772
Qy 599 -----ThrPheGlnPro-----ValAlaAlaGlyLeu 607
Db 5773 GCGATCCTCACTTCCACCTCGCGAGTCGCTGTTGATCAGATCATCGACGCGCGCATC 5832
Qy 608 -ProSer-----SerGlyAlaValGlyValMetPheHisAlaValProGlyLysG 624
Db 5833 GCGAGCTCGAGCACCTCGACATCGGCAACGCTTTGGCCTTGGGCAATGTGCGCGGGTG 5892
Qy 624 uGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsnGlyGlySerSe 644
Db 5893 AACCTCGTTTGGTAACTGCTGAGTTCACCTGGGTGCGGCAACGTCGCGCAACATC 5952
Qy 644 rTrpSerAlaIleThrGlyValSerSerAlaValAsnValGlyPheGlyLysSerAlaPr 664
Db 5953 AACGTCGGC-...CCGCAACCTCGGCGGAGCAACTGGGGTTGGCGACGTCGGGACC 6009
Qy 664 o-----GlySerSerTyrProAlaValPheValValGlyThrIleGlyG 679
Db 6010 GGCAACCTCGGTTGCGCAACATCGTTCGCGCAAT-TTCGGATTCGGCAACGCGGCT 6068
Qy 679 yValThrGlyAlaTyrArgSerAspCysGlyThrThrTrpValLeuIleAsnAspAs 699
Db 6069 GACCGCGCGCGCGGGGCTCGGCAATGTTGGG-...TTGGGTAAACGCCG 6116
Qy 699 pGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeuArgAr 719
Db 6117 CAGC-----GGCAGCTGGGGTTGGCCACGTCGGT---GTGGGCAATATCGGTT 6164
Qy 719 gValTyrIleGlyThrAsnGlyArgGlyIle---ValTyrGlyAsp----- 734
Db 6165 GGCCAAACCGGCAACGCGGATCGGATCGGCTGACCGGGGACTATCGGACCGGGAT 6224

Qy 734 eGlyGlyAlaProSerGly 740
Db 6225 CGCGCGCTGAACCTCGGC 6243

RESULT 3

US-11-136-527-2622
; Sequence 2622, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2622
; LENGTH: 7231
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-2622

Alignment Scores:
Pred. No.: 0.000712 Length: 7231
Score: 180.00 Matches: 201
Percent Similarity: 31.5% Conservative: 111
Best Local Similarity: 20.3% Mismatches: 320
Query Match: 4.5% Indels: 360
DB: 12 Gaps: 49

US-09-917-376-3 (1-740) x US-11-136-527-2622 (1-7231)

Qy 40 GlyMetTyrArg-----TrpAspAlaAlaAsnGlyArgTrpIleProLeu--- 54
Db 3121 GGAATATACAGAACTTCTGCTCTGGGACAGAAAT---CAGTGGACTTCTTCTGG 3177
Qy 55 -----LeuAspTrpValGlyTrpAsnAsnTrpGly----- 64
Db 3178 AGGAGACGATCAGAACTTCTACCTCTGGAATAGATGAGCAGCTGTCTTCTACTGG 3237
Qy 64 ----- 64
Db 3238 AAGAGAGGCTTAGAGAGCTGCTCTGGAGTAGAGACCTCAGTGGACTTCTTCTGG 3297
Qy 65 -----Tyr-AsnGlyValValSerIleAlaAlaAspProIleAs 77
Db 3298 AGAAGAGGTTCAAGAACTCTACTCTGGAATAGAGGACATAGTGTACTTCCGACTGG 3357
Qy 77 nThrAsnLysValTrpAlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspG 97
Db 3358 AGAAGTCCAGAACTTCTGCTCTGGAGTGGGAGACTTGAGTGGACTTCCCTCAGGAGG 3417
Qy 97 yAlaIleuArgSerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPhe 117
Db 3418 AGAAGTCTAGAAACATCTGCTTCAGGTGTAGAG---GATGTACCCAGCTTCTTACTGA 3474
Qy 117 sLeuGlyGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAs 137
Db 3475 AAGAGGAGGCTTAGAGACTTCTGCTCTGGAATAGAGGACATCACTGTCTTCTACTGG 3534
Qy 137 nAspAsnIleLeu-----TyrPheGlyAlaProSerG 148
Db 3535 AAGAGAAATCTAGAACTTCTGCTCTGGAGTAGAGGATGTCAGTGGACTTCTTCTGG 3594
Qy 148 yLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPhePr 168
Db 3595 AAAAGAGGCTTAGAGACTTCTGCTCTGGAATA-----GAGGACATTAGTGTCTCC 3648
Qy 168 oAsp-----ValGlyThrTyrIleAlaAsn---ProTh 178

Db 3649 TACTGACGACAGAGTCTAGAAACCTTCGCCCTCCGGGGATATGTCAGTGGGATTCCTTC 3708
Qy 178 rAspThrThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAspLysSe 198
Db 3709 TGGAGAGATGGTACAGAAACCTCTACTCTCTGGAGTACAGGGTGTGAGT-----GG 3759
Qy 198 rSerSerSerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspProAsnAs 218
Db 3760 TCTTCTCTGGAGAGAGGGTCTAGAAACGCTCTGCCTCTGGAGTAGAAGATCTTGGTCT 3819
Qy 218 nProValPheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProTh 238
Db 3820 TCCTACAGAGATAGTCTAGAGACTTCGCTCTCAGGAGTAGATGTACTGGGTATCCTTC 3879
Qy 238 rGly-----PheIleProHisLysGlyValPheAspProValAsnHi 252
Db 3880 TGGAGAGAGACACAGAAACCTCTGTCTCT-----GGGTAGGTAGTACCTTAGTGG 3933
Qy 252 sValLeuTyr-----IleAlaThrSerAsnThrGly----- 262
Db 3934 ACTTCTCTGGACAAAGAGCCTTAGACACCTCAGCTTCAGGAGCTGAGGACCTTGGTGG 3993
Qy 263 -----GlyProTyrAsp-----GlySerSerGlyAspValTrpLysPhe----- 275
Db 3994 CTGCTCTCTGGAAAGAGACCTTGGTAGGGTCTGCTTCTGGAGCCTCGGACTTCGGCAA 4053
Qy 276 -----SerValThrSerGlyThrTrpThrArgIleSerProValProSerTh 291
Db 4054 GCTACCTCTGGAACTCTAGAGAGTGGTCAAACTCCAGAACTAGTGGCCTTCCTCTCGG 4113
Qy 291 rAspThrAlaAsnAspTyrPheGlyTyrSerGlyLeu-ThrIleAspArgGlnHisProA 311
Db 4114 ATTTAGTGTGAA-----TATTCTGGAGTGGACATTTGGAAGTGGCCCATCTC 4161
Qy 311 snThrIleMetVal-----AlaThrGlnIleSer----- 320
Db 4162 TGGCCTGCTGACTTTAGTGGACTTCCATCCGGCTTCCCAACAGTCTCCCTTGTGGACAG 4221
Qy 321 -----Trp-----TrpProAsp----- 324
Db 4222 TACCTTAGTGAAGTATCATCAGCCACCATCCAGTGAACCTGGAAGGAGGGGACCAT 4281
Qy 325 -----ThrIlePheArgSerThrAspGly-----GlyAlaThrTrpThrArgIleT 341
Db 4282 CAGCGTAGTGGTTCAGGAGAAAGTACAGGCGGCCCTCAGTACGTGGAGTAGTGC 4341
Qy 341 rpAspTrpThrSerTyrProAsnArgSer-----LeuArgTyrValLeuA 356
Db 4342 G-GACATTAGTGGTCTCCCTTCAGAACTGAACCTCAGTGGGCCAAACATCCGGATCTCTG 4400
Qy 356 spIleSerAlaGluProTrpLeuThrPheGlyValGlnProAsnPro----- 371
Db 4401 ATGTCAGTGGAGAAACATCTGGATTTTGTATGTAGTGGACAGCCATTTGGGTCTTCTG 4460
Qy 372 -----ProValProSerProL 377
Db 4461 GCACGTGGTGAAGAACATCTGGGATTCCTGAAGTCAGTGGACAGGCCCTCAGGAGTCTG 4520
Qy 377 ysLeuGlyTyrMetAspGluAlaMetAlaIleAspProPheAsnSerAsp----- 393
Db 4521 ACACCACTGAGATATCTGAGCTTAGTGGACTGTCTCTGGGCAACCAAGATGTGAGTGAG 4580
Qy 394 -----ArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrL 411
Db 4581 AAGGCTCTGGAATCTCTTTGGCAGTGGCCCAATCTCTGCGCATAACATCT----- 4630
Qy 411 ysTrpAspSerGlyGlyGlnIleHisIleAlaPrometValLysGlyLeuGluThrA 431
Db 4631 -----GTGAGTGA-----GAAACCTCTG 4649
Qy 431 laValAsnAspLeuIleSerProProSerGlyAlaProLeuIleSerAlaLeu----- 448
Db 5633 erAlaVal-----AsnValGlyPheGlyLysSerAlaProGly-----SerSerTyrP 669
Db 5676 CTGCTCCCGTGAACACAGATCTGGAGATGACAAACCCAGGCGGAGGTGAGGACCAACC 5735
Db 4650 GGATTTCTGATCTCAGTGGGCAACCTTCAGGGTTCAGGTCTCAGGGGACACACCTG 4709
Qy 449 -----GlyAspLeuGlyGlyPheThrH 456
Db 4710 GAACCCCTGACCTGGCTTCTGGCGCATCAGTGGCAGTGGAGATCTTCTTGGCATTACGT 4769
Qy 456 isAlaAspValThrAlaVal-----ProSerThr----- 465
Db 4770 TTGTGGACACCACTCTATTGAAGTGACCCCTACCAACATTTAGAGAAGAGAGGGGTAG 4829
Qy 465 ----- 465
Db 4830 GGTCTGTGAACTCAGTGGCCTCCCTTCTGGGGAGACAGATCTGTCTGGCACATCTGGGA 4889
Qy 466 -----IlePheThrS 469
Db 4890 TGGTGGATCTCAGTGGACAACTCTTCTGGAGCAATCGATTCCAGTGGACTCATATCCCCA 4949
Qy 469 erProValPheThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIle----- 487
Db 4950 CTCCAGAAATTC---AGTGGCCTCCCAAGCGGAGTAGTCTGAGTCTGAGTGGAGAGGTCTCTG 5006
Qy 488 -----IleValArgAlaGlySerPheAsp----- 495
Db 5007 GAGTTGAGACTGGGAGCAGCTTGTCTCGGGAGCATTCGATGGCAGTGGACTTGTCTCAG 5066
Qy 496 -----ProSerSerGlnProAsnAspArgHisValAlaPheSer----- 508
Db 5067 GTTTCCTTACGGTGTCTCTGTAGACAGAACTTTGGTGGAGTCTATAACCTGGCTCCTA 5126
Qy 509 -----ThrAspGlyGlyLysAsn-----TrpPheGlnGlySerGluProG 522
Db 5127 CTGCTCAAGAACTCGAGAGGCGCTTCCAGCATTTTGGAAATTCAGTGGTGCCCATCTG 5186
Qy 522 lyGlyValThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheValTrpA 542
Db 5187 GAACACCAACATATCTGGGACCTCTCTGGGTCTTTAGACCAAGC-----ACATGGC 5240
Qy 542 laProGly-----AspProGlyGlnProValValTrpValAlaValGlyP 556
Db 5241 AGCTGGGTGGACAGAAAGCCAGACAGAGCCACCAAGCTCCCATATTTTAGTGGGACT 5300
Qy 556 heGlyAsnSerTrpAlaAlaSer----- 563
Db 5301 TCTTAGCACCACTCATGCAAGTGGAGATCCATACTGCCCCAACTGGCAGTGGGAAA 5360
Qy 564 --GlnGlyValProAlaAsnAlaGlnIleArgSerAspArgVal----- 577
Db 5361 CCTCGGGGCTTCCGGAAGTCACTTTAATCACTTCAGAGTTAGTGGAGGGGTGACTGAAC 5420
Qy 578 -----AsnProLysThrPheT 583
Db 5421 CAACTGTTTCCAGGAACCTGGCCATGGTCTTCTATGACATATCTCCCGGCTCTTTG 5480
Qy 583 yrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyValThrPheGlnPro- 602
Db 5481 AGCCAGTGGGGAAGCCCTCAGCATCTGGGGACCTTGGTGACCTGTAAACATCTTCCCG 5540
Qy 603 -----ValAlaAlaGlyLeuProSerSerGlyAla-----V 613
Db 5541 GGTCTGGGTAGAAAGCTTCAGTCCCAAGAGGAGCAGTACCCATCTGCTTACCTCGAGG 5600
Qy 613 alGlyValMetPheHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerS 633
Db 5601 CTGAGTGGGAGTGTCTGCTGCCCT-----GAGGCCAGCA 5636
Qy 633 erGlyLeuTyrHisSerThrAsnGlyLysSerSerTrpSerAlaIleThrGlyValSerS 653
Db 5637 GTCAATG-----TCTGAGTCCAGATCTGCATGTGAATCACCT 5675
Qy 653 erAlaVal-----AsnValGlyPheGlyLysSerAlaProGly-----SerSerTyrP 669
Db 5676 CTGCTCCCGTGAACACAGATCTGGAGATGACAAACCCAGGCGGAGGTGAGGACCAACC 5735

669	Qy	roAlaValPheValValGlyThr-----IleGlyGlyValT	681
5736	Db	CATGGACCTTTcAGGAAGGCACcAGGAGGAGTCCCGTGTCCAGAAGTcAGTGGAGAAT	5795
681	Qy	hrGlyAlaThrArgSerAspAspCysGlyThrThrTrpVal-----LeuI	696
5796	Db	CTAGCACTACTCCGACATAGATcAGcGACTTcAGGTGTGCCTTTTGCACGCCCATGA	5855
696	Qy	leAsnAspAspGlnHisGlnIThr---GlyAsnThrGlyGlnAlaIleThrGlyAspHis	715
5856	Db	CTTCtGGACACAGcGACTGAATcAGTGGAGAATGG-----TCtGATCATTA	5900
715	Qy	leAsnLeuArgArgValThrIleGlyThr	724
5901	Db	CCTcAGAGGTGAATGTTAcCGTcAGcACcC	5929

RESULT 4

```

US-11-052-554A-526
; Sequence 526, Application US/11052554A
; Publication No. US20050288666A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 526
; LENGTH: 9474
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-526

```

Alignment Scores:		
Pred. No.:	0.00133	9474
Score:	178.00	191
Percent Similarity:	33.4%	Conservative: 116
Best Local Similarity:	21.0%	Mismatches: 293
Query Match:	4.4%	Indels: 317
DB:	12	Gaps: 47

US-09-917-376-3 (1-740) x US-11-052-554A-526 (1-9474)

QY	43	ArgTrpAspAlaAlaAsnGlyArg- :::	-----Trp----- 	IlleProLeuLeuAspTrpValGlyTrpAsnAsnTrp :::::	----- 50
Db	521	CGTTGGGGGGTGGTTCGGCGCGTTCTCAATGCTCTATTTCGGACCCGCGAAGATCT			580
QY	51				63
Db	581	TCAGGCTTTAACCGCGGGCTTGGGCAATTCGGGTAAATTACAACGTCGGGTGGGCAATTCGG			640
QY	64	GlyTyrAsnGlyValValSerIleAlaAlaAspProIleAsnThrAsnIysValTrp----			82
Db	641	GGATAT -----TCAACCTGGCGCGACCAATGTCGGTCGCGAGAATTTGGGGT			688
QY	83	-----AlaAlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGly			97
Db	689	CTGCCAACCGCGTAGCGGGAAATTTCGGTTTCGGCAATATCGGCAACGCCAACTTCGGGT			748
QY	98	AlaIleLeuArg- :::			101
Db	749	TCGGCAACTCGGCTCTGGGTTCGCGCGGGCATGGGCAATATTGGTTGGGCAATCGGG			808
QY	102	-----SerSerAspGlnGlyAlaThrTrpGlnIleThrProLeuProPhoPheLys			117
Db	809	GCAGCAGCAACTACGGCTCGCAACCTGGGTCTGGGCAACATCTGGTTTGGTTCGCAACACGG			868

QY 361 -----ProTrpLeu 363
Db 1906 ATCAGATCGACCGCATCCCGCTGAACTCCGCGCCAGCGTCACTGTGCGCCCTATCTCTG 1965
QY 364 ThrPheGlyValGlnProAsnProProValProSerProLysLeuGlyTrpMetAspGlu 383
Db 1966 ATCAACGGGGTT-----AATATCCGGCG--ACCCCGGGCTTTGGCAACACGACCACC 2016
QY 384 AlaMetAlaLeuAspProPheAsnSer---AspArgMetLeu-----Tyr 397
Db 2017 GTCGCGTCTCGGGTTTCTCAACTCCGCGCAGCGTGGGTGTCGGGCTTCGGGAATTC 2076
QY 398 GlyThrGlyAlaThrLeuTyr-----AlaThrAsnAspLeuThrLysTrpAspSer 414
Db 2077 GGTGGCGGCACTCGGGTTGGTGGAAACAGCGGACAGCCGAGGTGGCTGGGGCGGGTTTCG 2136
QY 415 Gly-----GlyGlnIleHisIleAlaProMetValLys 425
Db 2137 GGTTCGCCAATTCGGTTCGCTGGGATCGGTGTGCTGAACCTTCGGCTCGGGTGTGTCG 2196
QY 426 GlyLeuGluGluThrAlaValAsnAspLeuIleSerProProSerGlyAlaProLeuIle 445
Db 2197 GGGCTGTACACACCGCGCGG-----TTGCCCGCGGGAGCCCGCGGTGTC 2244
QY 446 SerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThrAlaValProSerThr 465
Db 2245 TCGGCATCGGCAATGTTGGTCAG----- 2268
QY 466 IlePheThrSerProValPheThrThrGlyThrSerValAspTyrAlaGluLeuAsnPro 485
Db 2269 ---CAGCTGTCTGGGTTCTCGCGCGGGAGCGCA-----CTCAACACG 2310
QY 486 SerIleIleValArgAlaGlySerPheAspProSerSerGlnProAsnAspArgHisVal 505
Db 2311 AGCCTCATCACTCTCGGTTGGCCGATGTGGCAGCGTA-----AACGTC 2358
QY 506 AlaPhe-SerThrAspGlyGlyLysAsnTrpPheGln----- 517
Db 2359 GGTTCGGCAACGTCGGGACTTCAACCTGGGTGGCGCAATATCGCGACTTGAACGTG 2418
QY 518 ---GlySerGluProGlyGlyValThrThrGlyGlyThrValAlaAla-----SerAl 534
Db 2419 GGTTCGGCAATGTCTGGCGGGCGCAACGTCGGGTTCGGCAATATCGCGGATGCCAATTC 2478
QY 534 aAspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProValValTyrAlaVal 554
Db 2479 GGTTCGGCAATCGGGTCTCGCGCGGGCTCGCGCGG----- 2518
QY 554 IGlyPheGlyAsnSerTrpAlaAlaSer---GlnGlyValProAlaAsnAlaGlnIleAr 573
Db 2519 -----TGGGCAACATCGGGTTGGCAATGCGGACGCGCA----- 2554
QY 573 gSerAspArgValAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSe 593
Db 2555 -----ACGTCGGCTTCGGCAACATGGGTGTGGGC 2583
QY 593 rThrAspGlyGlyValThrPheGlnProValAlaAlaGlyLeu----- 607
Db 2584 AACATCGGGTTTCGGTAACACCGGCACCAACACCTCGGGATTTGGCTGACCGGGACAC 2643
QY 608 -----ProSerSerGlyAlaValGlyValMetPheHisAl 619
Db 2644 CAGACTGGGATCGGGGTTGAATCCCGT--GCCGGCAACATCGGG---TTGTTCAATC 2699
QY 619 aValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSer-----GlyLeuTyrHi 637
Db 2700 CGGCACCGGCAACGTCGGG---TTGTTCACTCCGGGACCGGGAATCTCGGGTTGTTCAA 2756
QY 637 sSer-----ThrAsnGlyGlySerSerTrpSerAlaIleThrGlyValSerSe 653
Db 2757 CTCGGGCACTTCAACACCGGCATCGGCAATGGGGAACGGGCGAGTACTGGGCTTTTCAA 2816
QY 653 rAlaValAsnValGlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheVa 673

Db 2817 TGCCGGTAATTTCAATACCGGTGTGCCCAACCTTGGTGTGTACACACCGGCGCTTCAA 2876
QY 673 lValGlyThrIleGlyGlyValThrGlyAlaTyrArgSerAspAspCysGlyThrThrTr 693
Db 2877 TGTGGGTGACCAAC-----ACCGTGTGTTTCAACCGCGGCGCATCAACACCGGCTG 2930
QY 693 pValleuIleAsnAspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAs 713
Db 2931 GTTCAACACCGGCAACGC-CAACACCGGCTGCCAATTTCGGGCAA----- 2975
QY 713 phisAlaAsnLeuArgArgValTyrIleGly 723
Db 2976 ----TGTGCACACCGCGCGCTCATGTGCGG 3002
RESULT 5
US-11-052-554A-527
; Sequence 527, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 527
; LENGTH: 7572
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-527
Alignment Scores:
Pred. No.: 0.00166 Length: 7572
Score: 175.50 Matches: 175
Percent Similarity: 33.3% Conservative: 92
Best Local Similarity: 21.8% Mismatches: 327
Query Match: 4.3% Indels: 209
DB: 12 Gaps: 38
US-09-917-376-3 (1-740) x US-11-052-554A-527 (1-7572)
QY 10 AsnValAlaIleGlyGlyGlyPheValAspGlyIleValPheAsnGluGlyAla--- 28
Db 4564 AATACCGGATCGTAAATTCGGGGATTGCCAGCACCGGGTTGTTCAACCGGGTGGGTTTC 4623
QY 29 ---ProGlyIleLeuTyrValArgThrAspIleGlyGlyMetTyrArgTrpAspAlaAla 47
Db 4624 AACACCGGTGTGTCATCGGTAGCTACACACCGCGAGTTTCAACCGCGGCGAGGCC 4683
QY 48 Asn---GlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyrAsn 66
Db 4684 AATACCGGCGGTTTCAACCGGCGAGTGTCAACACCGGCTGGTTGAACACCGGT----- 4737
QY 67 GlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaValGly 86
Db 4738 -----GACATCAACACCGGG----- 4752
QY 87 MetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSer---AspGln 105
Db 4753 ---GTGGCCAACTCGGCGAGCTCAACACCGGCGCTTCATCTCCGTAACAGCAAC 4809
QY 106 GlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGly--- 124
Db 4810 GCGCGCTTCTGGCGG-----GGCAGTACTACCGGCGCTG 4842
QY 125 ArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGly 144

```

Db      4843 CTGGCTTCTTACCCCGCGCTCTCCCAAA----- 4878
Qy      145 AlaProSerGlyLysGlyLeuThrArgSerThrAspSerGlyAlaThrTrpSerGlnMet 164
Db      4879 -----ACGCGCTTCTGGACCTC 4896
Qy      165 ThrAsnPheProaspValGlyThrTyriLeAlaAsnProThrAspThrThrGlyTyrGln 184
Db      4897 ACCCTCACCGCGGAGCTGGGCTCGCTGTTATCCCGCCCATCGACATTCGCGGATCCG 4956
Qy      185 SerAspIleGlnGlyValValTropVal--AlaPheAspLysSerSerSerLeuGly 203
Db      4957 CCGAGATTACGCCAACGTCGCATCGACAGCTTCACGTGCGGAGCATCCCGATTCCC 5016
Qy      204 Gln-----AlaSerLysThrIlePheValGlyValAlaAsp----- 215
Db      5017 CAGATCGACCTGCGCGCACACCGGTGAGCTCGGCTCGGCCCATCAGCTGCCGCAC 5076
Qy      216 -----ProAsnProValPheTrpSerArgAspGlyGlyAlaThrTrpGlnAla 232
Db      5077 CTCGATATTCCACGGGTGCGCGTCACGCTCAATTACTTGTGTTGCTCACAA----- 5127
Qy      233 ValProGlyAlaProThrGlyPheIleProHisLysGlyValPheAspProValAsnHis 252
Db      5128 ---CCGCGCGCGCCCTGAAATCGGTCGATTACGGGACTCTTCAACACCCCCATCGGC 5184
Qy      253 ValLeuTyriLeAlaThrSerAsnThrGlyGlyProTyriAspGlySerSerGlyAspVal 272
Db      5185 CTTACCCCTCGGGTGAACGACAGATGATCATCGTGCTAGCTCGTCGCAAGGACCATC 5244
Qy      273 TrpLysPheSerValThrSerGlyThrTrpThrArgIleSerProValProSerThrAsp 292
Db      5245 ACGGCTTCTCGCCAACTGCGCTTCAGCACCCCGTCTGCTCACCATTGACGAGATCCG 5304
Qy      293 ThrAlaAsnAspTyrPheGlyTyrSer-----GlyLeuThr 304
Db      5305 CTGCTGCCACGACATTACCGGCCACAGCGAGCCGCTCGACATCTTCCCGCGCGCTCACG 5364
Qy      305 IleAspArgGlnHisProAsnThrIle-----Met 314
Db      5365 ATCCCGCGATGAACCCGTCGACATCAACTGTCGGTGGCAGCGCGCGCTCACCAT 5424
Qy      315 ValAlaThrGlnIleSerTrpTrpAspThrIleIlePheArgSerThrAspGlyGly 334
Db      5425 CCGGCAATCACCATCGCGAAATCCCTTTGACCTCGTGCCACACGACGCTCGGC--- 5481
Qy      335 AlaThrTrpThrArgIleTrpAspTrpThrSerTyrProAsnArgSerLeuArgTyrVal 354
Db      5482 -----CCGTTTCATCCTC 5496
Qy      355 LeuAspIleSerAlaGluProTrpLeuThrPheGlyValGlnProAsnProProValPro 374
Db      5497 ATCGACTCTCCCGCGTCGCG-----GGGTTTCGG-----AATACGACCGGT 5538
Qy      375 SerProLysLeuGlyTrpMetAspGlu----- 383
Db      5539 GCTCCGTGTCGGGTTCTTCACTCGGTGCGGGTGGGTGTCGGGGTTGGGAATGTC 5598
Qy      384 ---AlaMetAlaIleAspProPheAsn---SerAspArgMetLeuTyrGlyThrGlyVal 401
Db      5599 GCGCGCATGTCGCGGTGCTGGAATCAGGCTCCGTCGCGGTGCTGGGTGGGGTTCG 5658
Qy      402 ThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlyGlnIleHisAla 421
Db      5659 GGTGTTTCAACCGCGGACGCTG-----CATTCGGGTGTCGTAATTCGCG 5706
Qy      422 ProMetValLysGlyLeuGluThrAlaValAsnAspLeuIleSerProProSerGly 441
Db      5707 TCTGGCATGTCGGGCTGTTCAACACGACGCGTGTGGGGTTC-----GGT 5751
Qy      442 AlaPro---LeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspValThr 460

```

```

Db      5752 CCGCGCGGTGGTGTGTCGGGTTTGGGTAGTGTGCGTCAG----- 5790
Qy      461 AlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSerValAspTyr 480
Db      5791 -----CAGTTCTCGGATGTTGGCGAGCGGACGCG----- 5823
Qy      481 AlaGluLeuAsnProSerIleValAlaArgAlaGlySerPheAspProSerSerGlnPro 500
Db      5824 -----CTGCATCAGGCTCGTCTCAATTTGGGTTGGCGCATGTGGGTG----- 5871
Qy      501 AsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGlySerGlu 520
Db      5872 -----GGCAATGTCGGTTCGCAATTCGGGGACTTTAACCTGGGTGCGGCAACGTC 5925
Qy      521 ProGlyGlyValThrThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheVal 540
Db      5926 GGTGGTTCATGTGGGTGGCGGCAATATCGCGGCAACATGTCGGGTGGGCAATGTC 5985
Qy      541 ---TrpAlaPro-----GlyAsp-----ProGlyGlnProValValTyr 552
Db      5986 GCGTGGGCAACTTTGGGCTCGGCAATTCGGGGTTAACCGCGGTCTGATGGTTGGGT 6045
Qy      553 AlaValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIle 572
Db      6046 AATATCGGTTTGTAAATCGCGGAGCTACAAATTCGGTTG---GCCAATATG----- 6096
Qy      573 ArgSerAspArgValAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPhe----- 590
Db      6097 -----GGTGTGGCAATATTCGGTTCCTAACACCGGCGATGGGAATTCGGTATT 6147
Qy      591 -----TyrArgSerThrAspGlyGlyVal 598
Db      6148 GGTGTGACCGGTGATATCTGACCGGTTTCGGTGTGTTTCAATACCGGAGCGGAATGTG 6207
Qy      599 ThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyVal---MetPhe 617
Db      6208 -----GGGTTGTTTAAATTCGGGACCGGTAATGTCGGGTCTCTTT 6246
Qy      618 HisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHis 637
Db      6247 AACTCTGCGACCGGAAAC-----TGGGGGTGTTCAATTTCGGGGAGTTAT--- 6291
Qy      638 SerThrAsnGlyLysSerTrpSerAlaIleThrGlyValSerSerAlaValAsnVal 657
Db      6292 AACACCGGATCGTAATTCGGGATTCGACGACGCGGTGTTTCAACGCGGGTGGGTTC 6351
Qy      658 GlyPheGlyLysSerAlaProGlySerSerTyrProAlaValPheValGlyThrIle 677
Db      6352 AACACGGGTGTGTCAATGCGGTAGCTACAAACCGGCGAGTTTCAACGCGGGCGAGGCC 6411
Qy      678 GlyGlyValThrGlyAlaTyrArgSerAspAspCysGlyThrThrTrpValLeuIleAsn 697
Db      6412 AAT-----ACGGCGGTTCACACCGGCGAGTGTCAACACGCGGTGTTGTGAACACCGGT 6465
Qy      698 AspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAla-----IleThrGly 712
Db      6466 GACATCAACCGGGGTGGCCAACTCCGCGACGTCACACCGGCGCTTCTATCTCCGCG 6525
Qy      713 AspHisAlaAsnLeuArgArgValTyrIleGlyThrAsnGlyArgGlyIleValTyrGly 732
Db      6526 AACTACAGCAAC-----GGCGCTTCTGCGCGGCGGCACTACAGGCG 6567
Qy      733 AsplleGly 735
Db      6568 CTGCTCGGC 6576

```

RESULT 6

```

US-11-052-554A-660
; Sequence 660, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE

```

; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
 ; FILE REFERENCE: 30853/40359A
 ; CURRENT APPLICATION NUMBER: US/11/052,554A
 ; CURRENT FILING DATE: 2005-02-07
 ; PRIOR APPLICATION NUMBER: US 60/589,227
 ; PRIOR FILING DATE: 2004-07-20
 ; PRIOR APPLICATION NUMBER: 1N 173/DEL/2004
 ; PRIOR FILING DATE: 2004-02-06
 ; NUMBER OF SEQ ID NOS: 763
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 660
 ; LENGTH: 15876
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli 0157:H7
 US-11-052-554A-660

Alignment Scores:			
Pred. No.:	0.012	Length:	15876
Score:	168.50	Matches:	191
Percent Similarity:	29.1%	Conservative:	83
Best Local Similarity:	20.3%	Mismatches:	270
Query Match:	4.2%	Indels:	337
DB:	12	Gaps:	47
US-09-917-376-3 (1-740) x US-11-052-554A-660 (1-15876)			
QY	27	GlyAlaProGlyLeuTyrValArgThrAspIleGlyMetTyrArgTrpAspAla	46
DB	12556	GGCGCTAATAGTTGCGCAAAATAGTACCGCAACGGCGCA	12597
QY	47	AlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyrAsn	66
DB	12598	-----ACCTGGGTGAAC-----	12609
QY	67	GlyValValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaValGly	86
DB	12610	-----GTGACCGTGGCGCAGACGCTGAC-----	12636
QY	87	MetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAspGlnGly	106
DB	12637	-----TGGAGTTACGTTGACGCGACGAACCTCACCAACGGCACC-----	12675
QY	107	AlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArgGly	126
DB	12676	ACCACCTGGCAGGTGGGTGTGTC-----GATCTGGCGGCAACGTT--GGCGCAACG	12726
QY	127	MetGlyGluArgLeuAlaValAspProAsnAsn-----	137
DB	12727	AGCAGCCAGTCGGGTGATCGATACCGTTAAACCCGGCGCAGGTGCTCACATGCCACG	12786
QY	138	-----AspAsnIleLeuTyr	142
DB	12787	ATCAGCACCACGACCGGGAGTTGCGCACTGACTTTATCACCAGCGCACCATGCTCAG	12846
QY	143	Phe-----GlyAlaProSerGlyLysGlyLeuTrpArgSerThrAsp	156
DB	12847	CTGACCGGTTCTGCTGGGCGCGGGCTGTCAGCGCGGAAGTGGCGCAGATTAGCCTTCAT	12906
QY	157	SerGlyAlaThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyrIleAlaAsn	176
DB	12907	AGCGCGCGACCTGGACACGCTGCCACCAACGGGTACACATGTGACATTACACCGACAGC	12966
QY	177	ProThrAspThrThrGly-----TyrGlnSerAspIleGlnGlyValValTrpVal	193
DB	12967	CGCAGCTGACCGACGGCAGCTACGTTTATCAGGTGCGGGTG-----	13008
QY	194	AlaPheAspLysSerSerSerLeuGlyGlnAlaSerLysThrIlePheValGlyVal	213
DB	13009	---CTGATCTGGCGGGGAACACCGCCCGCGTGGTGTCAAAACCGTGTGTCGATACG	13065
QY	214	AlaAspProAsn-----AsnProValPheTrpSerArgAspGlyGlyAlaThrTrp	230
DB	13066	ATTAAACCCACCGCCACCAACGATGTGTGTCATATGTCATCATGATGTGCGGCGCGCGAG	13126

```
QY 471 Val---PheThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleVal 489
Db 14008 ATCAGGTGCACCGTCGGCACCTGACGTGGTACGCGC-----TCGATTGTG--- 14052
QY 490 ArgAlaGlySerPheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThr 509
Db 14053 -----GCATTGTGAAGAGGCG----- 14070
QY 510 AspGlyGlyAsnTrpPheGlnGlySerGluProGly----- 522
Db 14071 GACGGCTATCTCGACTTCTGGATTGGTGCAGCTGCGCGCGGACTCCACACCTTCCTCG 14130
QY 523 -----GlyValThrGlyGlyThrValAlaAlaSerAlaAspGlySerArg 538
Db 14131 TGGAAACAGCAGCAGCAGCTGGTAGGCAACTCCACACGTCGACAGCGCGGTAGCGCC 14190
QY 539 PheValTrpAlaProGlyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsn 558
Db 14191 ACGTGGCGCGG-----GCGGTGACGGGTATCTT 14220
QY 559 SerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgVal--- 577
Db 14221 TCGCTCAAGGAAGTTCGTGGCGTCGATCTGACAAATGACGCGCAGGATCGACTGTTCCAG 14280
QY 578 -----AsnProLysThrPheTyrAlaLeuSer-----AsnGly 588
Db 14281 CACACCTATACCTCAACAACTATTACACGCTGCTTCGCTCATCAACACGAGGAATGGG 14340
QY 589 ThrPheTyrArgSerThrAspGlyGlyValThrPheGlnProValAlaAlaGlyLeuPro 608
Db 14341 ACGTTTGTCTGGGGCGAAGAACACCAACCAATACCTTCCTGACGGCGGCGGCGCT 14400
QY 609 SerSerGlyAlaValGlyValMetPheHisAlaValProGlyLys---GluGlyAspLeu 627
Db 14401 ATGAGCAGCAGCGTTCCATACCTACCTGGCCGATTCGATGTGTGACGGCATATGATCTC 14460
QY 628 TrpLeuAlaAlaSerSerGly-----LeuTyrHisSerThrAsnGly 641
Db 14461 TTCCTGCCCGCGCAGCAGGGAAGAGCTAACTACGCTCGCTGTTATTCAACACCAACGCG 14520
QY 641 ----- 641
Db 14521 GTACTGGGTGCCGGTGGCGGTGGCGCAACAGCAACCACTACGCCAGCCAGTTTACG 14580
QY 642 -----Gly 642
Db 14581 CTGGCGGTGGACTGGAACCAACGCGCTGATGGATATCGCCCGTATGCCCGCAGACCGGG 14640
QY 643 SerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnValGlyPheGlyLysSer 662
Db 14641 CAGTCGTAT---CTTTACTACTACGTACAGCAACGCCAGCAAC-----TGGACGCAATCG 14691
QY 663 AlaProGlySerSerTyrProAlaValPheValValGlyThrIleGlyGlyValThr--- 681
Db 14692 GCCCTCGCGCGCAGCCAGAGC-----GTTACCAACAGCGCGGTGGCGCA 14736
QY 682 -----GlyAlaTyrArg-----SerAspAspCysGlyThrThr 692
Db 14737 ATGACTACGACTGGAGCGCGCGGTGGATGTGTGTCTCCAAACAGTCGGCGGCGGTG 14796
QY 693 TrpValLeuIleAsnAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGly 712
Db 14797 TTCCTGAGCGCCACACCAACACCGGTAGCTACGGC-----ACTTCG 14838
QY 713 AspHisAlaAsnLeuArgValTyrIleGlyThrAsnGlyArgGlyIleValTyrGly 732
Db 14839 CTACACCTCGGCATC-----ACCGATCCCAACGCGATTAAAGTCTATTACGGC 14886
QY 733 Asp 733
Db 14887 AAT 14889
```

RESULT 7

```
US-11-052-554A-401
; Sequence 401, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 401
; LENGTH: 2943
; TYPE: DNA
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-401
```

```
Alignment Scores:
Pred. No.: 0.00326 Length: 2943
Score: 166.50 Matches: 164
Percent Similarity: 31.2% Conservative: 117
Best Local Similarity: 18.2% Mismatches: 302
Query Match: 4.1% Indels: 319
DB: 12 Gaps: 42
```

US-09-917-376-3 (1-740) x US-11-052-554A-401 (1-2943)

```
QY 57 TrpValGlyTyrAsnAsnTrpGly----- 64
Db 169 TATGTAGCTGGAACAACTATGCCCTGATAGTATTACAGCGGATATCTACCGGTTC 228
QY 65 -----TyrAsnGlyValValSer-----IleAlaAlaAspProIleAsn 77
Db 229 ATTAATAACGCTACCGTTAACGGCGTGATTCTACCTACTACTATCGACGCGGTATTCT 288
QY 78 ThrAsn-----LysValTrpAlaAlaValGlyMetTyrThr 89
Db 289 ACCAATACCAACGCAATAGTCTGACAATCAAAAACAGCACTATTTCACGGTATGATTAC 348
QY 90 AsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAspGlnGlyAlaThrTrp 109
Db 349 TCTGAGTCATGACTACTGATTGTGCT-----GATGACCGTCTACTGT 393
QY 110 GlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGlu 129
Db 394 TATGTTAT-----GAT 405
QY 130 ArgLeuAlaValAspProAsnAsn-----AspAsnIleLeuTyrPheGlyAla 145
Db 406 CGTCTGACACTGAGCGGTGATAATTCAACGATCGATGACAACTACGAGCATTA 459
QY 146 ProSerGlyLysGlyLeuTyrArgSerThrAspSerGlyAlaThrTrpSerGlnMetThr 165
Db 460 -----ACTTACAACGGTACTATAATAATGCCGCTGACACTCATGTGTGA 504
QY 166 AsnPheProAspValGlyThr----- 172
Db 505 GATGTTTACGATATGGGTACTGCTATTACCTGGATCAGGAAGTTGATGTGCCATCACT 564
QY 173 -----TyrIleAlaAsnProThrAspThrGlyTyrGlnSerAspIleGlnGly 189
Db 565 AATAACTCTCATGATGACGAGTATTACGCTGACTCAGGGGTATGAG----- 609
QY 190 ValValTrpValAlaPheAspLysSerSerSerLeuGlyGlnAlaSerLysThrIle 209
Db 610 -----TGGGAAGATATTGACGACAAACACAGTCAGCACTGGCGTAAACAGCAGCAAGTG 663
QY 210 PheValGlyValAlaAspProAsnAsnProVal----- 220
```

Db 664 TTT-----AATAACACTTATTACTGTTAAAGATTCTACTGTGACCTCT 705
Qy 221 ---PheTrpSerArgAspGlyGlyAlaThrTrpGlnAlaValProGlyAlaProThrGly 239
Db 706 GGTTCATGGACTGATGAAGT-----ACTACTGGT 735
Qy 240 PheIleProHisLysGlyValPheAspProValAsnHisValLeuTyr-----Ile 256
Db 736 TGGTTTGGCCATACTGTTAAATGCCAGCAACTATAGCAACACGCTGACTGCAGACGATGT 795
Qy 257 AlaThrSerAsnThrGlyGlyProTyrAspGlySerSerGlyAspValTrpLysPheSer 276
Db 796 GCAATTCCGCAATCGCAAAATCCGAT-----GCTGATAATGCGATCGACGACTACA 846
Qy 277 ValThrSerGlyThrTrpThrArgIleSerProValProSerThrAspThrAlaAsnAsp 296
Db 847 GTAACCTTAGCAACTCACTGATGGTGATGTTGTTTCTCCAGTAATTTGATGAA 906
Qy 297 TyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAla 316
Db 907 AACTTCTCCGCAAGTGTAAACAGCTATCGCGATGCTGATGGTGATGA----- 957
Qy 317 ThrGlnIleSerTrpTrpProAspThrIleIlePheArgSerThr---AspGlyGlyAla 335
Db 958 ---GATACCAACCGTTGGGTGGACACAGCCTGATGGATGTCACTCTGAACAACGCGCAG 1014
Qy 336 ThrTrpThrArgIle-----Trp 341
Db 1015 AAGTGGTGGCGTGCATCTCTGTTCAATGTTGATGAAGATGGTGATGTTCTTAC 1074
Qy 342 AspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGluPro 361
Db 1075 GACGATATGCTGTGTGACTGAAGCAACTGCACTCTGCTGATATTCGACTTAACAGC 1134
Qy 362 TrpLeuThrPheGlyValGlnProAsnProProValProSerProLysLeuGlyTrpMet 381
Db 1135 CTG-----TGGCCT 1143
Qy 382 AspGluAlaMetAlaIleAspProPheAsnSerAsp----- 393
Db 1144 TCATCAACTGTCGGTGTGTGATAACATCAATCAATATGACGAAATGGCCATATCGTA 1203
Qy 394 ---ArgMetLeuTyrGlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrp 412
Db 1204 GGAACGAAGTTTACGAGCGGTTGTTTAAATGTGACTTTGAACGGTGGTTCAGAGTGG 1263
Qy 413 Asp-----SerGlyGlyGlnIleHis 419
Db 1264 GATACACAAAATCTCTCTGATTGATATTTAAGTATTACAGCGGTTCCCAAGTTAAT 1323
Qy 420 IleAlaProMetValLysGlyLeuGluThrAlaValAsnAspLeuIleSerProPro 439
Db 1324 GTTGA-----GACTCTGCTGATCTCTGACACTGTCTCTGACT 1365
Qy 440 SerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyPheThrHisAlaAspVal 459
Db 1366 GCGCGTTCTAACCTG-----AACATCGGTGAAGACGGTCTATGACGACTAATACCCCTG 1419
Qy 460 ThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThrSerValAsp 479
Db 1420 ACCATGCACATAGTACCGTTAAATGTCGATGATGTTCTCGGGCTGGGGTTAGAA 1479
Qy 480 TyrAlaGluLeuAsnProSerIleIle-----ValArg 490
Db 1480 GATGTCGACTGTACGCAATACCATCACCGCTAACTACGACGGTCTGTTGGATATTAAAC 1539
Qy 491 AlaGlySerPheAspProSerSer---GlnProAsnAspArgHisValAlaPheSerThr 509
Db 1540 GTTGATCAGTTCGATGCTAAACCGTTCCAGGCCGATACCTGGAATCTGACCACTACCAC 1599
Qy 510 AspGlyGlyLysAsnTrpPheGlnGly----- 518

RESULT 8

US-11-052-554A-518

; Sequence 518, Application US/11052554A

Db 1600 GATACTAACGGCAACATTCAACGCTGGTGTATTTCGATATCCATAGCAGTGATTACGTAATG 1659
Qy 519 -----SerGluProGlyGlyValThrThrGly-----Gly 528
Db 1660 GATACCGATCTGTCACAGCATCGTACCAACGATACTACCAAGTCAAACTACGGTTATGGC 1719
Qy 529 ThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaProGlyAspProGlyGln 548
Db 1720 TTAATCGCAATGAACCTCTGATGCTACCTGACTATTAAACGGTAAACGGCGATAACGCAAC 1779
Qy 549 ProValValTyrAlaValGly-----PheGlyAsnSerTrpAlaAla 562
Db 1780 ACTGCTTCTATCGAAGCTGGTCAGAACGAAGTTGATAACAACGGTGACCATGTTGCAGCC 1839
Qy 563 SerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArg----- 576
Db 1840 GCGACCGGT-----AACTACAAAGTTGTTATCGACAACGCTACTCTGCTGCTTCT 1890
Qy 577 -----ValAsnProLysThrPheTyrAlaLeu 585
Db 1891 ATCGCTGACTACAACGGCAACGAGCTGCTACGTCAACGACAAAC----- 1938
Qy 586 SerAsnGlyThrPhe-----TyrArgSer 593
Db 1939 AGCAACGCGACCTTCTGCTGCTAAACAAAGCTGACCTGGTGCATACACATATCAGGT 1998
Qy 594 ThrAspGlyGlyValThr-----PheGlnProVal 603
Db 1999 GAACAGCGCGGTAAACACCGTTGTTCTGCAACAGATGGATTGACCGACTACGTTAAACATG 2058
Qy 604 AlaAlaGlyLeuProSerSerGly-----AlaValGly 614
Db 2059 GCGCTGAGCATCCCATCTCGGAACACCAATATCTGGAACCTGGAACAGACACCGTTGGT 2118
Qy 615 ValMetPhe-----HisAlaValProGlyLysGluGlyAspLeuTrpLeuAla 630
Db 2119 ACTCGTTTGACCAACTCTGCTCATGCGCTG---GCTGATAACGCGCGCATGGTAAAGC 2175
Qy 631 AlaSerSerGlyLeuTyrHisSerThrAsnGlyGlySerSerTrpSerAla---IleThr 649
Db 2176 TACTTCGGTGGTAACTTCAACGCGCAACCGCACCATCACTATGATCAGGATGTTAAC 2235
Qy 650 GlyValSerSerAlaValAsnValGlyPheGlyLysSerAlaProGlySerSerTrpPro 669
Db 2236 GGCATCATG-----GTGCGTGTGATACCAAAATTGACGGTAAACAC--- 2277
Qy 670 AlaValPheValValGlyThrIleGlyGlyValThrGlyAlaTyrArgSerAsp--- 687
Db 2278 GCTAAGTGGATCGTCGGTGGCTGCAGGCTTCGCTAAAGGTGACATGAATGACCGTTCT 2337
Qy 688 -----AspCysGlyThrThrTrpVal----- 694
Db 2338 GGTCAAGTGGATCAAGACAGCCAGACTGCTACATCTCTTCTGCTCATTTCGCGAAC 2397
Qy 695 ---LeuIleAsnAspAspGlnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAsp 713
Db 2398 AACGCTTTGTTGATGTTAGTGTAGCTACTCTCACTTCAACACGACCTGCTCGAACCC 2457
Qy 714 HisAlaAsn-----Leu 717
Db 2458 ATGACCAACGGTACTTACGTTGACGGTAGCACCAACTCCGACGCTTGGGGCTTCGGTTG 2517
Qy 718 ArgArgValTyr-----IleGlyThrAsnGlyArgGlyIleValTyrGlyAspIle 734
Db 2518 AAAGCGGTTACGACTTCAAACTGGGTGATGCTGGTTAGCTGACTCTCTTACCGCAGCAT 2577
Qy 735 GlyGly 736
Db 2578 TCTGGT 2583

Publication No. US20050288866A1

GENERAL INFORMATION:

APPLICANT: Sachdeva, et al.

TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE

FILE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL

FILE REFERENCE: 30853/40359A

CURRENT APPLICATION NUMBER: US/11/052,554A

PRIOR FILING DATE: 2005-02-07

PRIOR FILING DATE: 2004-07-20

PRIOR FILING DATE: 2004-02-06

NUMBER OF SEQ ID NOS: 763

SOFTWARE: PatentIn version 3.3

SEQ ID NO 518

LENGTH: 6615

TYPE: DNA

ORGANISM: Mycobacterium tuberculosis H37Rv

US-11-052-554A-518

Alignment Scores:

Pred. No.: 0.00946 Length: 6615
Score: 165.00 Matches: 196
Percent Similarity: 30.8% Conservative: 88
Best Local Similarity: 21.3% Mismatches: 292
Query Match: 4.1% Indels: 348
DB: 12 Gaps: 45

US-09-917-376-3 (1-740) x US-11-052-554A-518 (1-6615)

Qy 7 ThrTrpSerAsnValAlaLeuGlyGlyGlyPheValAspGlyLeuValPheAsnGlu 26
Db 5 ACCTGGTCTCC-ACAACGTCGGGAATGTCGGGCTTCTC-----AACGTC 48
Qy 27 GlyVala-----ProGlyLeuTyrValArgThrAspGlyGlyMetTyrArg-Tr 44
Db 49 GCGCGCTGGGATCGGGTGTGGGAATGTGGCAACACCATCTCGGGTATCTACAACGTG 108
Qy 44 pAspAlaAlaAsnGlyArgTrpIleProLeuLeuAspTrpValGlyTrpAsnAsnTrpGl 64
Db 109 GGCACGTGCGACCTCTGACGCGCGCGCTTAACCTCG-----GGTTGGCAA----- 154
Qy 64 yTyrAsnGlyValValSerIleAlaAspProfile-----AsnThrAs 79
Db 155 -----ATATCGGAACCAATATTATTCGCGCGCTGTCGCGACGCG 192
Qy 79 nLysVal-----TrpAlaAlaValGlyMetTyrThrAsnSerTrpAspPr 94
Db 193 GCGGGTACTCGCGCTATTAACTTTGGGCTTGGCCAAACACCGCAACCTCAACGTGGGCTTC 252
Qy 94 oAsnAspGlyAlaIleLeuArgSerAspGlnGlyAlaThrTrpGlnIleThrProLe 114
Db 253 GCAAGTCTCGCGGCTTTAACTTCG----- 277
Qy 114 uProPheLysLeuGlyGlyAsnMetProGlyArgGlyMetGlyGluArgLeuAlaValas 134
Db 278 -----GCGGCGCC 285
Qy 134 pProAsnAsnAspAsnIleLeuTyrPheGlyAlaPro---SerGlyLysGlyLeuTrpAr 153
Db 286 ACCATCGGCCACAAACAGTCGCGCATCGGGAACACCGGAATCTTCGATGTCGCGCTGGCG 345
Qy 153 gSer-----ThrAspSerGlyAlaThrTrpSerGlnMe 164
Db 346 AACCTGGGAGCTACAACTCGGCTTCGGAACCTTGGCGACGACCAACCTCGGCTTCGGC 405
Qy 164 tThrAsnPheProAspValGlyThrTrpIleAlaAsnProThrAspThrThrGlyTyrGl 184
Db 406 AACTTCG-----GCAGCTACACATCGGCTTCGGCAACGTTCGGCAACG----- 448
Qy 184 nSerAspIleGlnGlyValValTrpValAlaPheAspLysSerSer----- 199
Db 449 -----ACAATCTGGGTTTCGCTAACGCGGCGCGGCGCAACATCGGCTTT 492

Qy 200 -----SerSerLeuGlyGlnAlaSerLysThrIlePheValGl 212
Db 493 GCGAACACCGGACCAACAATGTCTGGTCTTGGAAACCGGACCAACAATGTCTGGCAATC 552
Qy 212 yValAlaAspProAsnAsnProValPheTrpSerArgAspGlyGlyAla----- 228
Db 553 GGGCTCACGGGCAACGGAC-----AGATCGGGTTCGGCAGCTTCAACTCG 597
Qy 229 -----ThrTrpGlnAlaValPr 234
Db 598 GGCAGCGGAACATCGGCTGTTCAACTCGGCGCAGCAACAACATCGGATTCTCAATTCC 657
Qy 234 oGlyAlaProThrGlyPheIleProHisLysGlyValPheAspProValAsnHisValLe 254
Db 658 GGCAGCGGCAACTTCGGC-ATCGCAAACTCGGCGAGCTTCAAC----- 699
Qy 254 uTyrIleAlaThrSerAsnThrGlyGlyProTyrAspGly-----SerSerGlyAspVa 272
Db 700 ----ACTGGCATCGGAACACCGGCAACACCAATACCGGCTTATTCACCTCGGCGAC-- 753
Qy 272 lTrpLysPheSerValThrSerGlyThrTrp-ThrArgIle-----SerProValProS 290
Db 754 -----GTCAACACGGGCGCTTCAACCGGCGAGCTTCAACACCGGTAGCTT 800
Qy 290 erThrAspThrAlaAsnAspTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisP 310
Db 801 CAACACCGGACGCTTCAACACCGGCGCT--TCAATCGGCGCAATACCAACACCGGCTA 857
Qy 310 roAsnThrIleMetValAlaThrGlnIleSerTrpTrpProAspThrIleIle-----P 328
Db 858 CCTCAACATTCGCAACTACAACACCGGATGCCAACACCGGCGAGCTTACACCGGGGC 917
Qy 328 heArgSerThrAspGlyGlyAlaThrTrp-----ThrArgIleTrpA 342
Db 918 TTTTCATCACCGGAACACTACAGCAACGGGTGTGTTTAAGCGCGAGTATCACCGGCTTG- 976
Qy 342 spTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGluProT 362
Db 977 -----TCGGCTCAACCT 989
Qy 362 rpLeuThrPheGlyValGlnPro-----A 370
Db 990 GGTGATCATGATGCTCTCCCAATAAGCTCGGCGTGAATATTCATCGATATCCGAT 1049
Qy 370 snProProValProSerProLysLeuGlyTrp----- 380
Db 1050 CACCGCTCGGCGCGCAACATCACCTTATGGCGCTCAGATTCCGCCACCGCGGATAT 1109
Qy 381 -----MetAspGluAlaMetAlaIleAspProPheAsnSerA 393
Db 1110 CGTCCTTTCATAGCGGCGCAGCGCCACTTTGGCGCCATTACATTCGGAACAT 1169
Qy 393 spArgMetLeuTyr--GlyThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTr 412
Db 1170 CACGGTTGTGCGGCCCCACGACGAGTCGCCATAGGAGGCGGATACCGCATCACCAT 1229
Qy 412 pAspSerGlyGlnIleHisIle-----AlaProMetValLysGlyLeuGl 428
Db 1230 AACTGGGCTGGCGCCATTAGGATCCCGCTCATCAGTATCCGCGGCGCGCAGGTTTCGG 1289
Qy 428 uGluThrAlaValAsn----- 433
Db 1290 AAATCGACCAACCAACCGCTCGTCAGGTTTCTTCAATACCGCGCGCGCGCTTCGGG 1349
Qy 434 -----AspLeuIleSerProProSerGl 441
Db 1350 CTTTCGGCAACTTCGGCGCGCAATTCGGGCTTTTGGAACTTGGCTTCGGACCTCGGG 1409
Qy 441 yAlaProLeuIleSerAlaLeuGlyAspLeuGly----- 452
Db 1410 GCGCTCGGGCTCTCAACGTCGCGCGCTCGGATCAGGTCTGGGGAACGTGGGCGCACAC 1469

```
QY 453 -----GlyPheThrHisAlaAspValThrAlaValProSerThrIlePheThrSerPr 470
Db 1470 CGTCTGGGGTCTTACACACAGCAGCTCGGACCTCGCGACCGCGGCTTCAATTACGG 1529
QY 470 oValPheThrThrGlyThrSerVal----- 478
Db 1530 CTGGCCCAACATCAGCACCATGATCGCGCGCTGCTGCGGACAGCAGCGGACCATGGT 1589
QY 479 -----AspTyrAlaGluLeuAsnProSerIleIleValArgAlaGl 492
Db 1590 CCTCAACCTGGGCTTGGCAACACCGGACCGCTCAACGTCGGCATT----- 1635
QY 492 ySerPheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGl 512
Db 1636 -----GCAAACTCGGCGATACAAACATCGGCTTTCGAAACCTCGGCAG 1679
QY 512 yLysAsnTrp-----PheGlnGlySerGluProGlyGlyValThrTh----- 526
Db 1680 CGCCAACTTCGGCAGCGCAATATCGGTGGCAACACATCGGCGCGCAACACCGGAAT 1739
QY 527 -----GlyGly-ThrValAlaAlaSerAlaAspGlySerA 538
Db 1740 ATTGACATCGGTTTGGCAATCTGGGAGTTTACAACTCGGCTTCGGAACTTCGGCGA 1799
QY 538 rgPheValTrpAlaProGlyAspProGlyGlnProValValThrAlaValGlyPheGlyA 558
Db 1800 TGACAACTCGGCTTC-GGAAACCTCGGCAGC-----TACAACTCGGCTTCGGAA 1849
QY 558 snSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValA 578
Db 1850 ACTTGGGCAACGACAACTCGGCTTC-----GCCAACACCGGC----- 1887
QY 578 snProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAspGlyGlyV 598
Db 1888 -----AGCAACATATCGGCTTCGGGACCGGCAGC----- 1920
QY 598 alThrPheGlnProValAlaGlyLeuProSerSerGlyAlaValGlyValMetPheH 618
Db 1921 -----AACAAATATCGGCTTCACGGGCGCGGCGGAGTCGG-----TTTCG 1966
QY 618 isAlaValProGlyLysGluGlyAspLeuTrpLeu-----AlaAlaSerSerG 634
Db 1967 GCTCCTGAAATCTGGCAGCGGAAACATCGGCTTGTTCAACTCGGCGCAGCGAAACATCG 2026
QY 634 lyLeuTyrHisSerThrAsnGly----- 641
Db 2027 GCTTTTCACTCGGCAACGGAACAGTTGGCATCGGCAACACCGGCAACCTTCG 2086
QY 642 -----GlySerSerTrpSerAlaIleThrGlyVal-----SerSerAlaValAsnValG 658
Db 2087 GGCTTGGAAACACCGGACGACCAACACCGGCTTCTTCACTCGGCGCAGCTCATACCG 2146
QY 658 lyPheGlyLysSerAla-----ProGlySerSerTyrProA 670
Db 2147 GTATCGGCAACACCGGAGCTTCAACACCGGAGCTTCAATCGGCGGATTCACACACCG 2206
QY 670 laValPhe-----ValValGlyThrIleGlyVal-----ThrG 682
Db 2207 GGGATTTCAACCCAGCGAGCTTACAAACACCGGATCGGAAACACCGGCGATGTTCACACG 2266
QY 682 lyAlaTyrArgSerAspAspCys-----GlyThrThrTrpValLeuIleAsnAspAspG 700
Db 2267 GGGCTTCATCTCGGAGCTACAGCAACGGGTCTTGTGG----- 2307
QY 700 lnHisGlnTyrGlyAsnTrpGlyGlnAlaIleThrGlyAspHisAlaAsnLeu 717
Db 2308 -----AGTGGAAATTTAT---CAGGGCCTCATTTGGCTTCGACGCGCGCTA 2349
```

RESULT 9

US-10-522-037-2/c

; Sequence 2, Application US/10522037

; Publication No. US20050282166A1

; GENERAL INFORMATION:

```
; APPLICANT: LIBRAGEN
; TITLE OF INVENTION: Method for the expression of unknown environmental DNA into adapted
; FILE REFERENCE: B0149WO
; CURRENT APPLICATION NUMBER: US/10/522,037
; CURRENT FILING DATE: 2005-01-24
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 37507
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: DNA sequence of clone FS3-135.
US-10-522-037-2
```

Alignment Scores:

Pred. No.:	0.107	Length:	37507
Score:	161.00	Matches:	181
Percent Similarity:	31.0%	Conservative:	67
Best Local Similarity:	22.6%	Mismatches:	277
Query Match:	4.0%	Indels:	276
DB:	8	Gaps:	44

US-09-917-376-3 (1-740) x US-10-522-037-2 (1-37507)

QY	36	ThrAspIleGlyGlyMetTyrArgTrpAspAlaAlaAsnGlyArgTrpIlePro-----	53
Db	26051	ACCGGACCTCGGCTCGCTCAGCTGG-----GCCAACGGCGAGACCCAGACC	25998
QY	54	-----LeuLeuAspTrpValGlyTrpAsnAsnTrpGlyTyrAsnGlyVal	68
Db	25997	ATCACGGTCCCGTGGCGAGCGACGGTATCGGTGAAGGCAACGAGACGTTCAACGTCGCG	25938
QY	69	ValSerIleAlaAlaAspProIleAsnThrAsnLysValTrpAlaAlaValGlyMetTyr	88
Db	25937	CTCTCCAATGCGACG-----GGTTCCACGGGTGTCTTC	25905
QY	89	ThrAsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAspGlnGlyAlaThr	108
Db	25904	ATCGGGCGCGCGGACCGCGACCGCATCATCTCTCG-----	25866
QY	109	TrpGlnIleThrProLeuProPheLysLeuGlyGlyAsnMetProGlyArgGlyMetGly	128
Db	25865	-----CCGTGGCCCGCGGTGGCAGCTGCTGCGGC-----	25833
QY	129	GluArgLeuAlaValAspProAsnAsnAspAsnIleLeuTyrPheGlyAlaProSerGly	148
Db	25832	-----TTCAAGACCTCGGGCGGC	25815
QY	149	LysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrpSerGlnMetThrAsnPhePro	168
Db	25814	GCCAAACGATGG---ACGGTCGGCTCCGACCGCTTCGAGGGCGCGCAGCCTTCGT	25758
QY	169	AspValGlyThrTyrIleAlaAsnProThrAspThrGlyTyrGlnSerAspIleGln	188
Db	25757	TCGGCGCAGGTGGTACGAGCGACCGACCACTCCAGCAGGTGAACCTCGACCTCTCG	25698
QY	189	-----GlyValValTrp-----ValAlaPheAspLysSerSerSerLeuGly	203
Db	25697	TTTCGGGGCAGCTGCTCGCGGCAACGCTCGCTTCGCCTATCGCTGTGAGCTTCCCC	25638
QY	204	GlnAlaSerLysThrIlePheValGlyValAlaAspProAsnAsnProValPheTrpSer	223
Db	25637	AACGCGCGCAAGTTCGAGTTCTCTG-----ATCGACGGCAGCGGTGTTTCAGC	25590
QY	224	ArgAspGlyGlyAlaThr-----TrpGlnAlaValProGlyAlaProThrGlyPhe	240
Db	25589	---GACGCGGGCACCACCGCTCGGCTCGAAGCTGCTGCTCCGGGTGACCGCGG--	25534
QY	241	IleProHisLysGlyVal-----PheAspProValAsnHisValLeuTyrIleAla-Th	258
Db	25533	CGCGCACACGCTGGTGTGGCGTTTCTCGAACTTCCTCTCCACGCTTGCACTCGCGAC	25474

```
Qy 258 rSerAsnThrGlyGlyProTyrAspGlySerSerGlyAsp---ValTrp----- 273
Db 25473 CTGGAACCCGCGCGCGCGATGCGCAACTGCGCGACCGCGCTGGATCGACGGCGT 25414
Qy 274 -LysPheSerValThrSerGlyThrThrThrArgIleSerProValProSerThrAspTh 293
Db 25413 GGCACCTCCCGGTTCGGGTGGGTGTCACCGCGCCCAACCAAGCGGTCTCTCG---GACCT 25357
Qy 293 rAlaAsnAspTyrPheGlyTyrSerGlyLeuThrIleAspArgGlnHisProAsnThrI 313
Db 25356 GAGCGCGAC-----GGCAGGAGCGACCTCTGCTG---CAGAACAGCGACGGCGCAT 25306
Qy 313 eMetValAlaThrGln----- 318
Db 25305 CGCGGTGTGACGATGAACGCGACCGGCATCACCAGCGCTGCGCAACCTGATGCGCGCGG 25246
Qy 319 -----IleSerTrpTrpProAsp-----ThrIleI 327
Db 25245 CGCGGGTGGACGGTGAAGCGAGCGCGCGACCTCAATGGCGACCTGCGCGCGCATCCT 25186
Qy 327 ePheArgSerThrAspGlyGlyAlaThrThrArgIleTrpAspTrpThrSerTyrPr 347
Db 25185 GTTCGCGCACCGATGGC----- 25167
Qy 347 oAsnArgSerLeuArgTyrValLeuAspIleSerAlaGluProTrpLeuThrPheGlyVa 367
Db 25166 ---CGCGCTACGCGTACATCATGACGCGGACTACG-----GTGACGCGCGCA 25120
Qy 367 lGlnProAsnProValProSerProLysLeuGlyTyrMetAspGluAlaMetAlaI 387
Db 25119 GGAGCTGCTGCC-----GCCGGCTGGCTGATCTGTAGCCACACGCGCG-- 25074
Qy 387 eAspProPheAsnSerAsp----- 393
Db 25073 ---GACCTGAACGCGAGCGAGCGAGCGACATCTCTGCGGCACACCGAGCGACGCGC 25018
Qy 394 -----ArgMetLeuTyr-----G1 398
Db 25017 GCACATCTGTTGATGACGCGCAGCATCATCGGCGCGCCAGCTCTTCGGGGCGG 24958
Qy 398 yThrGlyAlaThrLeuTyrAlaThrAsnAspLeuThr----- 410
Db 24957 GACCGGTGACGCTCACCAGCGCGCGACATGAACGCGCGACGCGACCTGGT 24898
Qy 411 -----LysTrpAspSerGlyGlyGlnIleHisIleAlaProMetValLysGlyLeuG1 428
Db 24897 CTTTCATGCACGCGCGCGCTACATCTACATC-----ATGACGCGCACCC 24847
Qy 428 uGluThrAlaValAsnAspLeuIleSerProProSerGlyAlaProLeuIleSerAlaLe 448
Db 24846 GGTGACGCGCGAGCGGCTTCTCTCGCGCGCGCGCTGG---ACGTAGCCACCT 24790
Qy 448 uGlyAspLeuGlyGly-----PheThrHisAlaAspValThrAl 461
Db 24789 GGGCGACCTCAACGCGCAGCGACGCGCGGACATGATCTTCGGCCACACCGCGCGCGC 24730
Qy 461 aValProSerThrIlePheThrSerProValPheThrThrGlyThrSerValAspTyrAl 481
Db 24729 G-----CACCTGTCTGTGATGAACGCGCACCATCTTCGGCGCCAGC-----GC 24688
Qy 481 aGluLeuAsnProSerIleIleValArgAlaGlySer-----Ph 494
Db 24687 CAGCTCTCTCCG-----GCCGCGCGGGTGGTGGTGAACCTGGTGGG 24643
Qy 494 eAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAs 514
Db 24642 CGACCTGAACGCGCGAGCGAGGACGACCTGCTTCCGCGCACACCGACGCGCGGTGCA 24583
Qy 514 nTrpPheGlnGlySerGluProGlyGlyValThrThrGlyGlyThrValAlaAlaSerAl 534
Db 24582 C-----ATCCGCGCTGAT 24571
```

```
Qy 534 aAspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProValValTyrAlaVa 554
Db 24570 GGATGGTGTTCATCATCATCCGCG---GGCGAT-----GCGCT 24535
Qy 554 lGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSe 574
Db 24534 GCCGCGGGGCGCGGTGGAAAGCTGACGAGCTGTGGAGCTGAACGGGACGACGCGCAA 24475
Qy 574 rAspArgValAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerTh 594
Db 24474 GGACCTGTGTC-----TTCCGCCACAC 24454
Qy 594 rAspGlyGlyValThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValG1 614
Db 24453 CGACGCGACGCTCACGGTGGC---CTGATGAACGGGCTCGCCACCATCGGACGCGCAA 24397
Qy 614 yValMetPheHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaSerSerG1 634
Db 24396 CTTGATC----- 24390
Qy 634 yLeuTyrHisSerThrAsnGlyGlySerSerTrpSerAlaIleThrGlyValSerSerAl 654
Db 24389 -----GCCCGCGCGGTGGTGGTC-----GC 24367
Qy 654 aValAsnValGlyPheGlyLysSerAlaProGlySerSerTyrPro-----AlaVa 671
Db 24366 ACCCGCGCAACCTCAGCTTGGCGGTTCCTCCACGGCGAGCGCCCGCGCTCGCGT 24307
Qy 671 lPheValValGlyThrIleGlyGlyValThrGlyAlaTyrArgSerAspCysGlyTh 691
Db 24306 TTTT-----CATGGCGCGCGGTATAGACTGGCGCGAAGACGTCC 24265
Qy 691 rThrTrpValLeuIleAsnAspAspGlnHisGlnTyrGlyAsnTrpGly-----G1 708
Db 24264 CTGGTGG-----TCCCATGATCTCCGACGACACGCGCTGGTTGGCGGGTTCTCTCG 24214
Qy 708 nAlaIleThrGlyAspHisAlaAsnLeuArgValTyrIleGlyThrAsnGlyArg 727
Db 24213 CGGTGCTGCTGGCGGCGACGCTGGTCTCGCTCGCTACGCGGTGCTGCAGCGCGC 24156
RESULT 10
US-10-793-626-3881
; Sequence 3881, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3881
; LENGTH: 3155
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3881
Alignment Scores:
Pred. No.: 0.019 Length: 3155
Score: 157.00 Matches: 168
Percent Similarity: 34.5% Conservative: 80
Best Local Similarity: 23.4% Mismatches: 299
Query Match: 3.9% Indels: 172
DB: 8 Gaps: 41
US-09-917-376-3 (1-740) x US-10-793-626-3881 (1-3155)
```


Db	959	CCA-----TATCTCGTAATTCACTACTGGTTGTTTGTGTACTTTTTCT	1003
Qy	367	ValGlnProAsnProValProSerProLysLeuGlyTrpMetAspGluAlaMetAla	386
Db	1004	GTTGGTTCACCTTCGCCAACATTTTCCCTGTTAAAGGGTCTTAGTTGGTGTTGA	1063
Qy	387	Ile-----AspProPheAsnSerAspArgMetLeuTyrglyThrGlyAla	401
Db	1064	ATTGTTTTTGTCTCTGGTTCACCTTTCTGTTTAACGCCTTTTACCTGGTTTAAATCA	1123
Qy	402	Thrlieu-----TyrAlaThrAsnAspLeuThrLysTrpAsp-----	413
Db	1124	GGATTGAATTCACGTTTCTTGTGCGAATGAATTTCTTCGTTGACGTGATCGGATCTCCA	1183
Qy	414	SerGlyGlyGlnIleHisIleAlaProMetValLysGlyLeuGluGluThrAlaValAsn	433
Db	1184	TCACTGGACCATAATTTTGCACATCATCACCACGTGGTGTGTACTCTCGCTGATACA	1243
Qy	434	AspLeuIleSerProSerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGly	453
Db	1244	GGGTTTTTAACCTCT-----GGTTTACCTGGAACTGCTCTTGGGTACCTTTCCGGTGCA	1297
Qy	454	PheThrHisAlaAspValThrAlaValProSerThrIlePheThrSerProValPheThr	473
Db	1298	TTT---CGATCAAATTCATCTCTATGTCCTGGCTTGATTTCTTCGCGCACCATAT----	1348
Qy	474	ThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAlaGlySer	493
Db	1349	-----TCTGTGAATTCATCTACTCTGGTTGT	1372
Qy	494	PheAspProSerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLys	513
Db	1373	TTT-----GTTATTTTCTGTTGGTTTACCTTCG	1402
Qy	514	AsnTrpPheGlnGlySerGluProGlyGlyValThrGlyGlyThrValAlaAspSer	533
Db	1403	CCAACTTTT-----TCCCCTGTTAAATGGGTTCTTAGTTGGTGTGTTAAATTGTTTT	1456
Qy	534	AlaAspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProValValTyrAla	553
Db	1457	GTTCCCTGGTTTCACCTTTTGGTTTAAACACGCTCTTCACCTGGT-----TTTAAA	1504
Qy	554	ValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArg	573
Db	1505	TCAGGATG---AATTCACGTTTCTTGTGCAATGGAAATTTCTTCGTTGACGTGATCGGA	1561
Qy	574	SerAspArgValAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSer	593
Db	1562	TCTCCATCAACTGGACCA-----TATTTTGTCAATCA	1594
Qy	594	ThrAspGlyGlyValThrPheGlnProValAlaAla-----GlyLeuProSer	609
Db	1595	TCCACAGGTGGAGTAATACTTCGCTGTATACAGGATTTTAAACCCCGGCTTACCT---	1651
Qy	610	SerGlyAlaValGlyValMetPheHisAlaValProGlyLysGluGlyAspLeuTrp---	628
Db	1652	---GGTTGGTTGTTGACHTACCTTCGGTGCATTTGGATCAAATTCATCCTTATGGCCT	1708
Qy	629	-----LeuAlaAspSerGlyLeuTyrglyThrHisSerThr	639
Db	1709	GGCTTGATTTCTTCGCCACCATAATGAACGATTTTCATCCACTGGTTGTTTGTACTTTTT	1768
Qy	640	AsnGlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnValGly---	658
Db	1769	TCTGTTGGTTTCACCTTCGCCAACTTTTCCCCCTGTTAAATGGGTTCTTAGTTGGTGT	1828
Qy	659	-----PheGlyLysSerAlaProGlySerSerTyrProAlaValPheValVal	674
Db	1829	GTAATTGTTTTT-----GTTCTGGTTTACCTTTTGAACGACTTCTCTGTA	1876

RESULT 11
US-10-793 -626--3107/c

```
; Sequence 3107, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3107
; LENGTH: 2733
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3107

Alignment Scores:
Pred. No.: 0.0215 Length: 2733
Score: 155.50 Matches: 164
Percent Similarity: 32.8% Conservative: 89
Best Local Similarity: 21.2% Mismatches: 287
Query Match: 3.9% Indels: 236
DB: 8 Gaps: 39

US-09-917-376-3 (1-740) x US-10-793-626-3107 (1-2733)
Qy 84 AlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSer 103
Db 2285 TCTGTGGTTCACCTTCGCCAACCTTTCTCTGTAATAGGTTCTTAGTT----- 2235
Qy 104 AspGlnGlyAlaThrTrpGlnIleThrPro---LeuProPheLysLeuGlyGlyAsnMet 122
Db 2234 ---GTTGGTGTGTAATGTTTGTTCCTGGTTTTCACCTTTTGAACGACTTCTCTGTA 2178
Qy 123 ProGlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsn---AspAsnIleLeu 141
Db 2177 CTTGGCGCTAAGTTGGATCAAAATCGCGTTTTTATCAACGGAATTTCTTCGTTGAC 2118
Qy 142 TyrPheGlyAlaProSer---GlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThr 160
Db 2117 GTAATCGATCTCCATCACTGACGACCATATTTTGTCAATCATCATCCACTGGTGTGACT 2058
Qy 161 TrpSerGlnMetThrAsnPhe-----ProAspValGlyThrTyrIleAla 175
Db 2057 ACTTCGCTGTATCAGGATTTTAACTCTGTTTACCTGGAACGCTCTCTTGGCTA--- 2001
Qy 176 AsnProThrAspThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPhe 195
Db 2000 ---CCITTCGGTTCGGATCAAAATTCAPCCTTA-----TGCGCTGGCTTG 1956
Qy 196 AspLysSerSer-----SerSerLeuGlyGlnAlaSerLysThrIlePheValGlyVal 213
Db 1955 ATTTCITTCGCCACCATATTCGTGATTTTCATCTACTGTTGTTGTTTATTTTCTGTT 1896
Qy 214 AlaAspProAsnAsnProValPheTrpSerArgAspGly-----GlyAlaThr 229
Db 1895 GGTTTCACCT---TGCCCAACTTTTCCCTGTTAATGGTTCTTAGTTGTTGGTGTGTA 1839
Qy 230 TrpGlnAlaValProGlyAlaPro-----ThrGlyPhe----- 240
Db 1838 ATTGTTTTTGTCTGTTTTCACCTTTCTGTTTAAACGCTCTTCACCTGGTTTAAAAATCA 1779
Qy 241 -----IleProHisLysGlyValPheAspProVal 250
Db 1778 GGATTGAATCAGTTTCTTGTGCAATGGAATTTCTTCGGTTGACGTAATCGGATCTCCA 1719
Qy 251 AsnHisValLeuTyrIleAlaThrSerAsnThrGlyGly
Db 263
```

```
Db 1718 TCAACTGGACCATATTTTGTACATCATCCACTGGTGTGACTACTTTCGCTGTATCA 1559
Qy 264 -----ProTyrAsp-----Gly 267
Db 1658 GGATTTTTTAACCTCTCGTGTAAAGTCTCTCTTGCTACTCTTTCGGTGCATTGGA 1599
Qy 268 SerSerGlyAspValTrpLysPheSerValThrSerGly----- 280
Db 1598 TCAAATTCATCTTATGGCTTGTGATTTCTTCGCCACCATATGAACGATTTTCATCT 1539
Qy 281 -----ThrTyrThrArgIleSerProValProSerThrThrAspThrAlaAsn 295
Db 1538 ACTGGTGTGTTTGTGTACTTTTCTGTTGTTTCTGCTTCCCAACTTTTCCCTGTAAAT 1479
Qy 296 AspTyrPhe----- 298
Db 1478 GGGTTCCTTAGTTGTGGTGTGTAATTTGTTTGTTCCTGGTTCACCTTTTGAACGACT 1419
Qy 299 -----GlyTyrSer 301
Db 1418 TTCTCTGTACCTGGCGCTAAGTTTGGATCAAAATTCGCGTTTTTATCAACGGAATTTCT 1359
Qy 302 GlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTrp 321
Db 1358 TCCGTTGACGT-GATCGGATCTCCATCACTGACCATATTTTGTACATCATCCACTGG 1300
Qy 322 TrpProAspThrIle---IlePheArgSerThrAspGlyGlyAlaThrTrp-ThrArgIle 340
Db 1299 TGGTGTGACTACTTCGCTGTATCAGGTTTTTAACCTCTGTTTACCTGGAACGCTCTC 1240
Qy 340 eTrp---AspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAl 359
Db 1239 TTGGTACTCTTTCGGTGCATTTGGATCAAAATCATCTTATGCGCTGGCTTGCATTTCTTC 1180
Qy 359 aGluPro-----TrpLeuThrPheGlyValGlnProAs 370
Db 1179 GCCACCATTAATGAACGATTTTCATCCACTGGTGTGTTTGTACTTTCTCTGTTGTTT 1120
Qy 370 nProProValProSerProLysLeuGlyTrpMetAspGluAlaMetAlaIle----- 387
Db 1119 TTGCGCAACTTTTCCCTGTTAATGGTTTCTTAGTTGTGTTGTGTTGTTTGTGTTTGT 1060
Qy 388 -----AspProPheAsnSerAspArgMetLeuTyrGlyThr-----GlyAlaThr 402
Db 1059 TCCTGGTTTACCTTTTGAACGACTTTCTCTGTACTCTGGCGCTAAGTTTGGATCAAAATTC 1000
Qy 402 rLeuTyrAlaThrAsnAspLeuThrLysTrpAsp-----SerGlyGlyGly 417
Db 999 GCGTTTTTATCAACGGAATTTCTTCGCTGACGTAATCGGATCTCCATCACTGACCC 940
Qy 417 nIleHisIleAlaProMetValLysGlyLeuGluGluThrAlaValAsnAspLeuIleSe 437
Db 939 ATATTTTGTACATCATCCACTGGTGGAGTAACACTTTCGCTGTATCAGGATTTTAAAC 880
Qy 437 rPro-----ProSerGlyAla----- 442
Db 879 TCCTGGTTTACCTGGAACGCTCTCTTGGCTACTCTTTCGGTGCATTTGGATCAAAATTCATC 820
Qy 443 -----ProLeuIleSerAlaLeuGlyAspLeuGlyGlyPh 454
Db 819 CTTATGGCTGGCTTGAATTTCTTCGCCACCATATTTCTGTGATTTTCATCTACTGTTGTTT 760
Qy 454 eThrHisAlaAspValThrAlaValProSerThrIlePheThrSerProVal----- 471
Db 759 TGTACTTTTCTGTGTTGTTCA---CCTTCGCCAACTTTT---TCCCTGTGTTAATGGGTT 706
Qy 472 -PheThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleIleValArgAl 491
Db 705 CTTAGTTGTGGTGTGTAATTTGTTTGTCTGCTGGTTTTCACCTTTCTCTGTTAACCGCTC 646
Qy 491 a-----GlySerPh 494
Db 645 TTTACCTGGTTTTAAATCAGGATTCAAATTCAGTTTCTTGTGCAATGGAATTTCTTCCT 586
```

QY 494 eAsp-----ProSerSerGlnProAsnAspArgHisValAlaPheSerThrAs 510
DB 585 TGACGTGATCGGATCTCCATCACTGGACCA----- 555
QY 510 pGlyGlyLysAsnTrpPheGlnGlySerGluProGlyValThrThrGlyGlyThrVa 530
DB 554 -----TATTTTGTGCATCATCCACTGGTGTGTGACTACTTCGCCTGTA-- 510
QY 530 lAlaAlaSerAlaAspGlySerArgPheValTrpAlaProGlyAspProGlyGlnProVa 550
DB 509 -----TCAGGGTITTTA--ACTCTCTGGTTTACTTGAACG----- 477
QY 550 lValTrpAlaValGlyPheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAl 570
DB 476 -----TCTCTTGGCTACCTTCCTGGTGCATTTGGATCAAAATTC 439
QY 570 aGlnIleArgSerAspArgValAsnProLysThrPheTrpAlaLeuSerAsn--GlyTh 589
DB 438 ATCCTTATGGCCTGGCTGTGATTTCTTCCGCCACCATATTTCTGTGATTTCACTACTGTTG 379
QY 589 rPheTrpArgSerThrAspGly-----GlyValThrPheGlnProValAlaAlaGlyLe 607
DB 378 TTTTGTATTTTCTGTGTGGTTTCACTTCCCAACTTTTCCCTGTATATGGTCTT 319
QY 607 uProSerSerGlyAlaValGlyValMetPheHisAlaValProGlyLysGluGlyAspLe 627
DB 318 A-----GTTGTGTGTGTGTGTAATTTTGTCTCTGGT-----TCACCTTT 277
QY 627 uTrpLeuAlaAlaSerSerGlyLeuTrpHisSerThrAsnGlyLysSerTrpSerAl 647
DB 276 TTGTTTAAACAGCTCTTCACTGGTTTAAATCA-----GGATTGAATTCACGTTTCTT 223
QY 647 aIleThrGlyValSerAlaValAsnValGlyPheGlyLysSerAlaProGlySerSe 667
DB 222 GTCGAATGAATTTCTTCCTGGTGTGATCGGATCTCCATCACTGACCA----- 171
QY 667 rTyProAlaValValGlyValGlyThrIleGlyGlyValThr 681
DB 170 -----TATTTTGTGCATCATCCACAGGTGGAGTAAT 138

RESULT 12

US-11-136-527-3719
; Sequence 3719, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3719
; LENGTH: 2595
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-3719

Alignment Scores:
Pred. No.: 0.0224 Length: 2595
Score: 155.00 Matches: 163
Percent Similarity: 29.3% Conservative: 55
Best Local Similarity: 21.9% Mismatches: 276
Query Match: 3.8% Indels: 252
DB: 12 Gaps: 35

US-09-917-376-3 (1-740) x US-11-136-527-3719 (1-2595)

QY 12 AlalleGlyGlyGlyPheValAspGlyIleValPheAsnGluGlyAla----- 28

DB 454 GCTGTGGTGTATTGGTGGCATCGGTGGCTAGGAGTCTCAACAGGTGCTGTGGTCCT 513
QY 29 -----ProGlyIleLeuTrp 33
DB 514 CAACCTGGAGCTGGAGTCGGACCGGAGAAAGCCTGGGAAAGTTCTCTGTGTGC----- 567
QY 34 ValArgThrAspIleGlyMetTrp--ArgTrpAspAlaAlaAsnGlyArg----- 50
DB 568 -----GGTCTTCCAGGTGTATACCAGGTGGAGTGTCTCCAGGAAACAGGAGTCGG 618
QY 51 -----TrpIleProLeuLeuAspTrp-----V 58
DB 619 TTCCTCTGTGTGGAGTGTCTCTCGAGTTCCTCCACCTGGCACAGGAGTCAAGGCCAAGTT 678
QY 58 aLGIYTrpAsnAsnTrp-GlyTrpAsnGlyValValSerIleAla----- 72
DB 679 CCGGTGGAGGAGGTGTCTTTTCTGGAATCCAGGGTGGGCCCTTTTGGGGTTCAG 738
QY 73 -----AlaAspProIleAsnThrAsnLysValTrpAlaAlaValGly 86
DB 739 CAGCCTGTGTCCACTGGTGTATCCCATCAAGCCCAAGCTGCCAGGTGGCTATGGA 798
QY 87 Met---TyThrAsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSerAspGln 105
DB 799 CTGCCCTATACCAATGG----- 816
QY 106 GlyAlaThrTrpGlnIleThrProLeuProPheLysLeuGlyGly----- 120
DB 817 -----AACTGTCCTATGAGTGTGTGTGGTGGAGGGGCAAGCT 855
QY 121 AsnMetPro---GlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsn 139
DB 856 GGTATCCCAACAGGACAGGAGGTGGTCCAGGACGACGTGGCAGCAGCTAAA---GCA 912
QY 140 IleLeuTrpPheGlyAlaProSerGlyLysGlyLeuTrpArgSerThrAspSerGlyAla 159
DB 913 GCGAAGTATGCAAGTGTCT--GGAGGAGGTGGAGTCTCTCCCTGTGTGGAGGGGTGGC 969
QY 160 ThrTrpSerGlnMetThrAsnPheProAspValGlyThrTyrrIleAlaAsnProThrAsp 179
DB 970 ATTCCTGTGTGTGTGGCGCAATCTCTGGATTTGA----- 1005
QY 180 ThrThrGlyTyrrGlnSerAspIleGlnGlyValValTrpValAlaPheAspLysSerSer 199
DB 1006 -----GGCATTCAGCGCTGGAACTCCCGCAGCAGCTGTCTGTGCA 1047
QY 200 SerSerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspProAsnAsnPro 219
DB 1048 AAGGTGCGCGCTAAGGCTGCTAAGTATGGAGCTGTGGAGGCTTAGTGCCTGTGGACCA 1107
QY 220 ValPheTrpSerArgAspGlyAlaThrTrpGlnAlaValProGlyAlaProThrGly 239
DB 1108 GGAGTTAGGTTCCCAAGTGTCTGGAATCCAGAGTGTGGATCCCAAGT-----GTCCGT 1161
QY 240 PheIleProHisLysGlyValPheAspProValAsnHisValLeuTyrrIleAlaThrSer 259
DB 1162 GGCATTCAGGTTGGTGGCATTCAGGTGTGGTGGCATC-----CCA 1206
QY 260 AsnThrGlyGlyProTyrrAspGlySerSerGlyAspValTrpLysPheSerValThrSer 279
DB 1207 GGTGTGGGGGCTGTGATCGAGGTCCAGGATTTGTG----- 1245
QY 280 GlyThrTrpThrArgIleSerProValProSerThrAspThrAlaAsnAspTyrrPheGly 299
DB 1246 GGTGGACCAAGGGGCTGTGTACAGCTGCAGCTGCCAAAGTGTCTGCCAAAGTGCACAG 1305
QY 300 TyrSerGlyLeuThrIleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIle 319
DB 1306 TATGGA-----GCCAG-AGGTGG 1322
QY 320 SerTrpTrpProAspThrIleIlePheArgSerThrAspGlyAlaThrTrpThrArg 339

Db 1323 AGTTGGCATCCGAC-----ATACGGGTTGGTGGTGGCTTCTCGGCTATGG 1373
 Qy 340 IleTrrpAspTrpThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAla 359
 Db 1374 TGTGGAGCTGGAGC-----TGGAC-TTGGAGGTCAAGCC 1408
 Qy 360 GluProTrpLeuThrPheGlyValGlnProAsnProValProSerProLeuGly 379
 Db 1409 AAGCTGCTGCTG-----CCGCTGCCGCTGCCAAGCTG--- 1441
 Qy 380 TrpMetAspGluAlaMetAlaIleAspPropheAsnSerAspArgMetLeuTyrGlyThr 399
 Db 1441 ----- 1441
 Qy 400 GlyAlaThrLeuTyrAlaThrAsnAspLeuThrLysTrpAspSerGlyGlnIleHis 419
 Db 1442 -----CCAAAGTACGGTCCCGAGGAGCTGCAACCTGGGAGGCTGGTGCAGGTGCA- 1494
 Qy 420 IleAlaProMetValLysGlyLeuGluThrAlaValAsnAspLeuIleSerProPro 439
 Db 1495 GTACCAAGTGCACCTGCCAGGTGCAGTACCAGGTGCACGT-----CCA 1536
 Qy 440 SerGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspVal 459
 Db 1537 GGTGCAGTACCA-----GGTGCAGTGCAGGT----- 1563
 Qy 460 ThrAlaValProSerThrIlePheThrSerProValPheThrThrGlyThr---SerVal 478
 Db 1564 ---GCAGTGCAGGTGTGCCGGGCACTGGTGGAGTGCAGGAGCAGGTACCCCTCGCGCT 1620
 Qy 479 AspTyrAlaGluLeuAsnProSerIleValAlaGlySerPhe-----AspPro 496
 Db 1621 GCAGCTGCTCCCGCGCTGCCAGGAGCGCCCAAGCGCAGCAGTATGTTGGGCGCT 1680
 Qy 497 SerSerGlnProAsnAspArgHisValAlaPheSerThrAspGlyGlyLysAsnTrpPhe 516
 Db 1681 GGT-----GTGGTGGCGCTTCCCGGTGGAGTTGGT----- 1710
 Qy 517 GlnGlySerGluProGlyGlyValThrThrGlyGly---ThrValAlaAlaSerAlaAsp 535
 Db 1711 GTTGTGGGCTTCCCGTGGAGTTGGCCCTGGTGGTGTACTGTATGGAACTGGTCCT 1770
 Qy 536 GlySerArgPheValTrpAlaProGlyAspProGlyGlnProValValTyrAlaValGly 555
 Db 1771 GGCACCGCGCTTGTA-----CCTGTGTATCTTGA-----GGA 1803
 Qy 556 PheGlyAsnSerTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAsp 575
 Db 1804 GCAGGGACACACAGCTGCTGCCAAATCTGCTGCTAAAGCGCGCTGCCAAAGCCAG- 1857
 Qy 576 ArgValAsnProLysThrPheTyrAlaLeuSerAsnGlyThrPheTyrArgSerThrAsp 595
 Db 1858 -----TACAGCTGCTGCT 1872
 Qy 596 Gly-----GlyValThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAla 612
 Db 1873 GGCCTTGGGCGCGTGTCCCTGGACTGGGTGGTGTGGTGTCCCGGATTTGGGGCT 1932
 Qy 613 -----ValGlyValMetPheHisAlaValPro 621
 Db 1933 GTGCTGCTGGATTTGGGCGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1992
 Qy 622 GlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsnGly 641
 Db 1993 GGATCC-----CTGGCTGCATCCAAAGCTGCTAAATACGAGCAGCAGGT 2037
 Qy 642 GlySerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnValGlyPheGlyLys 661
 Db 2038 GGCCTTGGAGGCGCTGGAGGACTTGGTGGCGCTGGAGGACTTGGTGGCGCTGGAGGATTC 2097
 Qy 662 SerAlaProGlySerSerTyrProAlaValPheValValGlyThrIleGlyGlyValThr 681
 Db 2098 GGTGGGCGCTGGAGT-----CTCGGTGGTGTCTCCTGGTGGAGTAGCA 2139

Qy 682 Gly 682
 Db 2140 GGA 2142
 RESULT 13
 US-10-793-626-3661
 ; Sequence 3661, Application US/10793626
 ; Publication No. US20050255478A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KIMMERLY, WILLIAM JOHN
 ; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
 ; FILE REFERENCE: PUI480US
 ; CURRENT APPLICATION NUMBER: US/10/793,626
 ; CURRENT FILING DATE: 2004-03-04
 ; PRIOR APPLICATION NUMBER: 60/164,258
 ; PRIOR FILING DATE: 1999-11-09
 ; NUMBER OF SEQ ID NOS: 4472
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3661
 ; LENGTH: 3050
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic
 ; OTHER INFORMATION: nucleic acid sequence
 US-10-793-626-3661
 Alignment Scores:
 Pred. No.: 0.0375 Length: 3050
 Score: 153.00 Matches: 172
 Percent Similarity: 33.0% Conservative: 95
 Best Local Similarity: 21.3% Mismatches: 295
 Query Match: 3.8% Indels: 248
 DB: 8 Gaps: 41
 US-09-917-376-3 (1-740) x US-10-793-626-3661 (1-3050)
 Qy 84 AlaValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaIleLeuArgSerSer 103
 Db 11 TCTGTGTGTTCACCTTCCCAACTTTTCCCTGTAATGGTCTCTAGTT----- 61
 Qy 104 AspGlnGlyAlaThrTrpGlnIleThrPro---LeuProPheLysLeuGlyGlyAsnMet 122
 Db 62 --GTGTGTGTGTAATTTGTTTCTCTGTTTCACTTCTGTTTAAACGCGCTCTTAA 118
 Qy 123 ProGly-----ArgGlyMetGlyGluArgLeuAlaValAspProAsnAsnAspAsnIle 140
 Db 119 CCTGGTTTTAAATCAGGATTCAGTTTC---TTGTGCAATGGAATTTCTTCCGTT 175
 Qy 141 LeuTyrPheGlyAlaProSer---GlyLysGlyLeuTrpArgSerThrAspSerGlyAla 159
 Db 176 GACGTGATCGGATCTCCATCACTGGACCAATTTTGTGCATCATCATCATCGTGTGTG 235
 Qy 160 ThrTrpSerGlnMetThrAsnPhe-----ProAspValGlyThrTyrIle 174
 Db 236 ACTACTTGCCTGTATCAGGCTTTTAACTCCTGCTGTACCTGGACGCTCTCTGGCTA 295
 Qy 175 AlaAsnProThrAspThrThrGlyTyrGlnSerAspIleGlnGlyValTrpValAla 194
 Db 296 -----CCTTTCGGTGCATTTGGATCAAATTCATCTTA-----TGGCTGCGC 337
 Qy 195 PheAspLysSerSer-----SerSerLeuGlyGlnAlaSerLysThrIlePheValGly 212
 Db 338 TTGATTTCTTTCGCCACCATATTCCTGTGATTTCTACTGTTGTTTGTATTATTTCT 397
 Qy 213 ValAlaAspProAsnAsnProValPheTrpSerArgAspGly-----GlyAla 228
 Db 398 GTTGGTTCACCT---TCGCCAACTTTTTCCTGTTAATGGTCTCTTAGTGTGTGT 454
 Qy 229 ThrTrpGlnAlaValProGlyAlaPro-----ThrGlyPhe--- 240
 Db 455 GTAATTGTTTGTCTCTGTTTACCTTTTGTGTTTAAACGCTCTCTACCTGTTTAA 514

QY 241 -----IleProHisLysGlyValPheAspPro 249
DB 515 TCAGGATTGAATTCACGTTCTTGTGCGAATGAATTTCTCGTTGACGTGATCGGATCT 574
QY 250 ValAsnHisValLeuTyrIleAlaThrSerAsnThrGlyGly
DB 575 CCATCACTGGACCAATATTTGTGCATCATCCACAGGTGGAGTAACACTTTCGCTGTGA 634
QY 264 -----ProTyrAsp----- 266
DB 635 TCAGGATTTTAAACCCCGCTTACCTGTTGCGTTGTTGACATACCTTTCGGTGCAATT 694
QY 267 GlySerSerGlyAspValTyrLysPheSerValThrSerGly- 280
DB 695 GGATCAAAATTCATCCTATGCGCTGTGATTTCTTCGCCACCAATAATGAACGATTTC 754
QY 281 -----ThrTyrThrArgIleSerProValProSerThrAspThrAla 294
DB 755 TCCACTGGTTGTTGTACTTTTCTGTTGCTGCTTCACTTCGCCCAACTTTTCCCTGTT 814
QY 295 AsnAspTyrPhe----- 298
DB 815 AATGGTTCTTAGTTGTTGTTGTTGTAATGTTTGTCTGCTGTTACCTTTTGAACG 874
QY 299 -----GlyTyr 300
DB 875 ACTTTCTCTGACCTGGCGTAAGTTTGGATCAAAATTCGCGTTTATCAAAACGGAATT 934
QY 301 SerGlyLeuThrIleAspArgLysHisProAsnThrIleMetValAlaThrGlnIleSer 320
DB 935 TCTCCGTTGACGTAAT-CGGATCTCCATCACTGGACCAATATTTGTGCATCATCCAC 993
QY 321 TrpTyrProAspThrIle---IlePheArgSerThrAspGlyGlyAlaThrTyr-ThrAr 339
DB 994 TGGTGGATACTACTTCGCTGTATCAGGTTTAACTCCTGCTTACCTGGACGTC 1053
QY 339 gileTyr---AspTyrThrSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSe 358
DB 1054 CTCCTGGTACTCTTCGCTGATTTGGATCAAAATTCATCCTTATGCGCTGCTGATTC 1113
QY 358 rAlaGluPro-----TrpLeuThrPheGlyValGlnPr 369
DB 1114 TTCGCCACCAATAATGAACGATTTTCATCCACTGTTGTTGTTGTTGTTGTTGTTGTT 1173
QY 369 oAsnProValProSerProLysLeuGlyTyrMetAspGluAlaMetAlaIle----- 387
DB 1174 ACCTTCGCCAACTTTTCCCTGTTAATGGTTCTTAGTTGTTGTTGTTGTTGTTGTTT 1233
QY 388 -----AspProPheAsnSerAspArgMetLeuTyrGlyThr-----GlyAl 401
DB 1234 TGTTCTCGTTACCTTTTGAACGACTTCTCTGACTCGCGCTAAGTTTGGATCAAA 1293
QY 401 aThrLeuTyrAlaThrAsnAspLeuThrLysTyrAsp-----SerGlyGly 416
DB 1294 TTCGGGTTTTTATCAACGGAATTTCTTCGCTGGACGTAATCGGATCTCCATCACTGG 1353
QY 416 yGlnIleHisIleAlaProMetValLysGlyLeuGluThrAlaValAsnAspLeuI 436
DB 1354 ACCATATTTGTCATCATCCACTGCTGGTGAAGTAACACTTTCGCTGATCAGGATTTT 1413
QY 436 eSerPro-----ProSerGlyAla----- 442
DB 1414 AACTCTGGTTTACCTGGAACGCTCTCTTGGCTACCTTTCGGTGCAATTTGGATCAAAATTC 1473
QY 443 -----ProLeuIleSerAlaLeuGlyAspLeuGlyGly 453
DB 1474 ATCCTTATGGCTGGCTGATTTCTTCGCCACCAATTTCTGTGATTTCTACTGTTGTTG 1533
QY 453 yPheThrHisAlaAspValThrAlaValProSerThrIlePheThrSerProVal----- 471
DB 1534 TTTTGTACTTTTCTGTGTTGTTCA---CCTTCGCCAACTTTT---TCCCTGTTAATGG 1587

QY 472 ----PheThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSerIleValar 490
DB 1588 GTTCTTAGTTGTTGGTTGTAATTTGTTCTGCTGTTACCTTTCTGTTTACGCG 1647
QY 490 gAla-----GlySe 493
DB 1648 CTCCTTACCTGGTTTTAAATCAGGATTGAATTCACGTTTCTGTCGAATGGAATTTCTTC 1707
QY 493 rPheAsp-----ProSerSerGlnProAsnAspArgHisAlaValaPheSerTh 509
DB 1708 CGTTGACGRTGATCGGATCTCCATCACTGGACCA----- 1741
QY 509 rAspGlyGlyLysAsnThrPheGlnGlySerGluProGlyGlyValThrThrGlyGlyTh 529
DB 1742 -----TATTTTGTACATCATCCACTGCTGGTGTGACTACTTCGCTGT 1785
QY 529 rValAlaAlaSerAlaAspGlySerArgPheValTyrAlaProGlyAspProGlyGlnPr 549
DB 1786 A-----TCAGGGTTTTTA---ACTCCTGGTTTACCTGGAACG-- 1819
QY 549 oValValTyrAlaValGlyPheGlyAsnSerTyrAlaAlaSerGlnGlyValProAlaAs 569
DB 1820 -----TCCTCTTGGCTACCTTTCGGTGCATTTTGGATCAAA 1854
QY 569 nAlaGlnIleArgSerAspArgValaAsnProLysThrPheTyrAlaLeuSerAsn--G1 588
DB 1855 TTCACTCTTATGGCTGCTGCTGATTTCTTCGCCACCAATATTTCTGATTTTCACTACTGG 1914
QY 588 yThrPheTyrArgSerThrAspGly-----GlyValThrPheGlnProValAlaAlaG1 606
DB 1915 TTGTTTGTGTTATTTTCTGTTGTTTCACTTCGCCCACTTTTTCCTGTTTAAATGGGTT 1974
QY 606 yLeuProSerSerGlyAlaValGlyValMetPheHisAlaValProGlyLysGluGlyAs 626
DB 1975 CTTA-----GTTGTTGGTGTGTAATTTGTTTGTCTCTGTT-----TCACC 2016
QY 626 pleuTyrLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsnGlyGlySerSerTrpSe 646
DB 2017 TTTTGTGTTTAAACAGCTCTTCACTGTTTAAATCA-----GGATTGAATTCACGTTT 2070
QY 646 rAlaIleThrGlyValSerSerAlaValaAsnValGlyPheGlyLysSerAlaProGlySe 666
DB 2071 CTTGTGCAATGAATTTCTTCGCTGACGTGATCGGATCTCCATCACTCACTGACCA----- 2125
QY 666 rSerTyrProAlaValPheValValGlyThrIleGlyGlyValThrGlyAlaTyrArgSe 686
DB 2126 -----TATTTGTCAATCATCCACAGGTGAGTAACACTTTCGCTGTATC 2172
QY 686 rAsp-----AspCysGlyThrThrTyrValLeuIleAs 697
DB 2173 AGGATTTTAAACCCCGCTTACCTGTTGCTGTTGTTGACTACCTTTCGGTGCATTTG 2232
QY 697 nAspAspGlnHisGlnTyrGly 704
DB 2233 -GATCAAAATTCATCTTATGGC 2253
RESULT 14
US-10-793-626-3970/c
; Sequence 3970, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3970
; LENGTH: 3748
; TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: synthetic

OTHER INFORMATION: nucleic acid sequence

US-10-793-626-3970

Alignment Scores:

Pred. No.: 0.0549 Length: 3748
Score: 152.00 Matches: 162
Percent Similarity: 33.0% Conservative: 91
Best Local Similarity: 21.1% Mismatches: 294
Query Match: 3.8% Indels: 220
DB: 8 Gaps: 38

US-09-917-376-3 (1-740) x US-10-793-626-3970 (1-3748)

Qy 85 ValGlyMetTyrThrAsnSerTrpAspProAsnAspGlyAlaLeuLeuArgSerSerAsp 104
Db 2516 GTTGGTTCACTTCACCAACTTTTCCCTGTTAATGGTTCTTAGTT----- 2469
Qy 105 GlnGlyAlaThrTrpGlnIleThrPro---LeuProPheLysLeuGlyGlyAsnMetPro 123
Db 2468 GTTGGCGTTGTAATGTTTGTTCCTGGTTACCTTTTGAACGACTTTCTCTGTACCT 2409
Qy 124 GlyArgGlyMetGlyGluArgLeuAlaValAspProAsnAsn---AspAsnIleLeuTyr 142
Db 2408 GGGCGTAAGTTGGATCAAAATTCGCGTTTATCAACGGAATTTCTTCGTTGACGTA 2349
Qy 143 PheGlyAlaProSer---GlyLysGlyLeuTrpArgSerThrAspSerGlyAlaThrTrp 161
Db 2348 ATCGGATCTCATCACTGACCAATATTTGTTCATCATCATCTGCTGGTGGTGGTACT 2289
Qy 162 SerGlnMetThrAsnPhe---ProAspVal---GlyThrTyrIleAlaAsnProThr 178
Db 2288 TCGCCTGTATCAGGATTTTAACTCTCTGGTACTGTTGCGTGTGTGTACTACCTTTC 2229
Qy 179 AspThrThrGlyTyrGlnSerAspIleGlnGlyValValTrpValAlaPheAspLysSer 198
Db 2228 GGTGCGTTCGGATCAAAATTCATCCTTA-----TGGCCTGCTTGAATTTCTTCG 2181
Qy 199 SerSer-----SerLeuGlyGlnAlaSerLysThrIlePheValGlyValAlaAspPro 216
Db 2180 CCACATAATGACGATCTCATCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2121
Qy 217 AsnAsnProValPheTrpSerArgAspGly-----GlyAlaThrTrpGlnAla 232
Db 2120 ---TCGCCAACTTTTCTCTGTTAATGGTCTTAGTTGTTGTTGTTGTTGTTGTTT 2064
Qy 233 ValProGlyAlaPro-----ThrGlyPheIleProHisLysGlyValPheAspProVal 250
Db 2063 GTTCTGTTTCACTTTTGAACGACTTCTCTGTACCTGCGCTAAGTTTGGATCAAT 2004
Qy 251 AsnHisValLeu----- 254
Db 2003 TCGCGTTTATCAACGGAATTTCTTCGTTGACGTAATTCGGATCTCCATCACTGGA 1944
Qy 255 ---TyrIleAlaThrSerAsnThrGlyGly----- 263
Db 1943 CCAATTTTGTACATCATCATCACTGTTGGTGTGACTACTTCCGCTGTATCAGGATTTTA 1884
Qy 264 -----ProTyrAsp-----GlySerSerGly 270
Db 1883 ACTCTGGTTTACTCGAAGCTCTCTCTGGTACTTTTCGGTGGTTCGGATCAATTC 1824
Qy 271 AspValTrp-----LysPheSerValThrSerGlyThrTrpThr 283
Db 1823 TCCTTATGGCCTGGCTTGAATTTCTTCGCCACCATATCTGTGATTTCTACTCTGGTGT 1764
Qy 284 ArgIle-----SerProValProSerThrAspThrAlaAsnAspTyrPhe 298
Db 1763 TTTGTTATTTTCTGTGTTCACTTCGCAACTTTTCCCTGTGTTAATGGTCTCTTA 1704
Qy 298 ----- 298

Db 1703 GTTGTGGTGTGTAATGTTTGTTCCTGTTACCTTTCTGTTTAACACGCTCTTCA 1644
Qy 299 -----GlyTyrSerGlyLeuThr 304
Db 1643 CCTGGTTTAAATCAGGATTGAATTCAGTTTCTTGTGCAATGGAATTTCTTCCGTTGAC 1584
Qy 305 IleAspArgGlnHisProAsnThrIleMetValAlaThrGlnIleSerTrpProAsp 324
Db 1583 GTAAT-CGGATCTCCATCACTGACCATATTTTGTACATCATCACTGGTGGTGTGAC 1525
Qy 325 ThrIle---IlePheArgSerThrAspGlyAlaThrTrp-ThrArgIleTrp---As 342
Db 1524 TACTTCGCTGTATCAGGATTTTAACTCCTGGTTTACCTGGAAGCTCTCTTGGCTACC 1465
Qy 342 pTrpSerTyrProAsnArgSerLeuArgTyrValLeuAspIleSerAlaGluPro-- 361
Db 1464 TTTCCGTCATTTGGATCAAAATTCATCTTATGCGCTGCTGTGATTTCTTCCACCAT 1405
Qy 362 -----TrpLeuThrPheGlyValGlnProAsnProProVa 373
Db 1404 ATGAACGATTTTCATCTACTGTTGTTTGTGTACTTTTCTGTGTGTTACCTTCGCAAC 1345
Qy 373 lProSerProLysLeuGlyTrpMetAspGluAlaMetAlaIle-----As 388
Db 1344 TTTTCCCTCTTAAATGGTTCCTTAGTTTGTGGTGTGTGTAATTTTGTTCCTGGTTC 1285
Qy 388 pProPheAsnSerAspArgMetLeuTyrGlyThr-----GlyAlaThrLeuTyrAl 405
Db 1284 ACCTTTGAACGACTTCTCTGACCTGGCGCTGAAGTTGGATCAAAATTCGCTTTT 1225
Qy 405 aThrAsnAspLeuThrLysTrpAsp-----SerGlyGlyGlnIleHisI 420
Db 1224 ATCAACGGAATTTCTTCGTTGACGTATCGGATCTCCATCACTGGACCATATTTGT 1165
Qy 420 eAlaProMetValLysGlyLeuGluThrAlaValAsnAspLeuIleSerProSe 440
Db 1164 CACATCATCCACTGGTGGTGTGACTACTTCGCTGTATCAGGTTTAACTCTCT- 1110
Qy 440 xGlyAlaProLeuIleSerAlaLeuGlyAspLeuGlyGlyPheThrHisAlaAspValTh 460
Db 1109 -GGTTTACCTGGAACGCTCTCTGGCTACCTTTCGGTGCATTT---GGATCAAAATTCATC 1054
Qy 460 xAlaValProSerThrIlePheThrSerPro-----ValPheThrThrGlyThrSe 477
Db 1053 CTTATGCGCTGGCTTGAATTTCTTCGCCACCATATGAACGATTTTCATCGCTGGTGT 994
Qy 477 xValAspTyrAlaGluLeuAsnProSer----- 486
Db 993 TGTTCATTTTCTGTGTGTTTCACCTTCGCCAACCTTTTCCCTGTTAATGGTTCCTAGT 934
Qy 487 -----IleIleValArgAlaGlySerPheAspProSerSerGlnProAsnAspAr 503
Db 933 TGTGGTGTGTAATGTT-----TTTGTTCCTGGTTCA-----CCTTT 895
Qy 503 gHisValAlaPheSerThrAspGlyGlyLys----- 513
Db 894 TTGAACGACTTCTCTGTACCTCGGCTAAGTTTGGATCAAAATTCGCGTTTTATCAAA 835
Qy 514 -----AsnTrpPheGlnGlySe 519
Db 834 CGGAATTTCTTCGTTGACGTAATCGGATCTCCATCACTGGACCATATTTTGTACATC 775
Qy 519 xGluProGlyGlyValThrThrGlyThrValAlaAlaSerAlaAspGlySerArgPh 539
Db 774 ATCCACTGGTGGAGTAACACTACTTCGCTGTA-----TCAGGAT 736
Qy 539 eValTrpAlaProGlyAspProGlyGlnProValValTyrAlaValGlyPheGlyAsnSe 559
Db 735 TTTA---ACTCTGGTTTACCTGGAACG-----TCCTC 706
Qy 559 rTrpAlaAlaSerGlnGlyValProAlaAsnAlaGlnIleArgSerAspArgValAsnPr 579
Db 579 ----- 579

705	TTGGCTACCTTTTCGGTGCATTTGGATCAAAATTCATCTTATGGCGCTGGCTTGATTCTTTC	646
579	oLysThrPheTyAlaLeuSerAsn---GlyThrPheTyArgSerThrAepGly----	596
645	GCACCAATATCTGTGATTTCACTACTGGTGTGTTGTTACTTTTCTGTGGTTCA	586
597	-GlyValThrPheGlnProValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMe	616
585	TTCCGCAACTTTTTCCTCTGTTAATGGGTCTTA-----GTTGTTGGTGTGTGT	538
616	pHeHisAlaValProGlyLysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTy	636
537	AATGTGTTTGTTCCTGGTTCACT-----TTCGTTTAACGGCGCTCTTTACC	490
636	rHisSerThrAsnGlyGlySerSerTrpSerAlaIleThrGlyValSerSerAlaValAs	656
489	TGTTTTTAATCAGGATTGAATTCACGTTCTTGTGCAATGGAAATTTCTTCGGTTGACGT	430
656	nValGlyPheGlyLysSerAlaProGlySerSerTyProAlaValPheValValGlyTh	676
429	GATCGGATCTCCATCACTGGACCA-----TATTTTGTGCATCATCATC	388
676	rIleGlyGlyValThr	681
387	CACTGGTGGTGTGACT	372

RESULT 15
US-10-793-626-4165/c
Sequence 4165, Application US/10793626
Publication No. US20050255478A1
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/10/793,626
CURRENT FILING DATE: 2004-03-04
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1993-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4165
LENGTH: 3600
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4165

Alignment Scores:		
Pred. No.:	0.0577	Length: 3600
Score:	151.50	Matches: 183
Percent Similarity:	34.6%	Conservative: 91
Best Local Similarity:	23.1%	Mismatches: 303
Query Match:	3.8%	Indels: 215
DB:	8	Gaps: 46
US-09-917-376-3 (1-740) x US-10-793-626-4165 (1-3600)		
QY	27	GlyAlaPro-----GlyIleLeuTyrValArgThrAspIleGlyMetTyrArgTrp 44
Db	3600	GGATCTCCATCACTGGACCATATTTTGTGCACATCATCCACTGGTGGTG---ACTACT 3544
QY	45	AspAlaAlaAsnGlyArgTrpIleProLeuLeu-----AspTrpValGlyTrp 60
Db	3543	TCGCCTGTATCAGGGTTTTTAACCTCTGGTTTACCTCGAACCTCCTCTTGCTACCTTTC 3488
QY	61	AsnAsnTrpGlyTyrAsn-----GlyValValSerIleAlaAlaAspPro 75
Db	3483	GGTGCAATTGGATCAAAATTCATCCTTATGGCTGGCTTGATTTCTTCGCCACCATATCT 3420
QY	76	IleAsnTrpAsnLys-----ValTrpAlaAlaValGlyMetTyrThrAsnSerTrp 92

D	b		3423	G T G A T T T C A T C T A C T G G T G T T T G T T A T T T T T T C T G T T G G T T C A C C T T G C C A C A C T T T T	3366
Q	y		93	A s p P r o A s n A s p G l y A l a I l e L e u A r g S e r S e r A s p G l i n G l y A l a T h r T r p G l i n I l e T h r	112
D	b		3363	T C C C C T G T T A A T G G G T C T T A G T T - - - - - G T T G G T G T T A A T T G T T T T T G T T	3316
Q	y		113	P r o - - - L e u P r o P h e L y s L e u G l y G l y A s n M e t P r o G l y - - - - - A r g G l y M e t G l y G l u	129
D	b		3315	C C T G G T T C A C C T T T T G T T T A A C A C G C T C T T C A C C T G G T T T T A A A T C A G G A T T G A A T T C A	3256
Q	y		130	A r g L e u A l a V a l A s p P r o A s n A s n A s p A s n I l e L e u T y r P h e G l y A l a P r o S e r - - - G l y	148
D	b		3255	C G T T T C - - - T T G T C G A A T G G A A T T T C T T C C G T T G A C G F G A T C G G A T C C C A T C A A C T G G A	3199
Q	y		149	L y s G l y L e u T r p A r g S e r T h r A s p S e r G l y A l a T h r T r p S e r G l i n M e t T h r A s n P h e - - -	167
D	b		3198	C C A T A T T T T G T C A C A T C A T C C A C A G G T G G A G T A A C T A C T T C G C C T G T A T C A G G A T T T T A	3139
Q	y		168	- - - P r o A s p V a l - - - G l y T h r T y r I l e A l a A s n P r o T h r A s p T h r T h r G l y T y r G l i n S e r	185
D	b		3138	A C C C C G G C T T A C C T G G T G C G T T G T T G A C T A C C T T T C G G T G C A T T T G G A T C A A A T T C A	3079
Q	y		186	A s p I l e G l i n G l y V a l V a l T r p V a l A l a P h e A s p L y s S e r S e r - - - - - S e r	200
D	b		3078	T C C T T A - - - - - T G S C C T G G C T T G A T T T C T T G C C A C C A T A A T G A A C G A T T T C A	3031
Q	y		201	S e r L e u G l y G l i n A l a S e r L y s T h r I l e P h e V a l G l y V a l A l a A s p P r o A s n - - - A s n P r o	219
D	b		3030	T C A C T G G T - - - T G T T T G T T A C T T T T T C T G T T G T T C A C C T T G C C A A C T T T T T C C C C T	2974
Q	y		220	V a l P h e T r p S e r A r g A s p G l y A l a T h r T r p G l i n A l a P r o G l y A l a P r o - - - - -	237
D	b		2973	G T T A A T G G G T C T T A G T T G T T G T T G T A A T T G T T T T T C T C T G G T T C A C C T T T T T T G A	2914
Q	y		238	T h r G l y P h e I l e P r o H i s L y s G l y V a l P h e A s p P r o V a l A s n A s i s V a l L e u T y r I l e A l a	257
D	b		2913	A C G A C T T T C T C T A C C T G C G C T A A C T T T T G G A T C A A A T T C G C G T T T T T T A T C A A A C G G A	2854
Q	y		258	T h r S e r A s n T h r - - - - - G l y G l y P r o T y r A s p G l y S e r S e r G l y A s p V a l T r p L y s	274
D	b		2853	A T T C T T C C G T G A C G T A A T C G G A T C T C C A - - - - - T C A C T G G A C C A T A T T T T G T C	2803
Q	y		275	P h e S e r V a l T h r S e r G l y T h r T r p T r p A r g I l e S e r P r o V a l P r o S e r T h r A s p T h r A l a	294
D	b		2802	A C A T C A T C A C C T G G T G A G T A A C T A C T - - - - - T C G C C T G T A T C A G G - - - - -	2761
Q	y		295	A s n A s p T y r P h e G l y T y r S e r G l y L e u T h r I l e A s p A r g G l i n H i s P r o A s n T h r I l e M e t	314
D	b		2760	- - - - - T T T T T A A C T C C T G T T T A - - - - - C C T G G A A C G - - - - -	2734
Q	y		315	V a l A l a T h r G l i n I l e S e r T r p T r p P r o A s p T h r I l e P h e A r g S e r T h r A s p G l y G l y	334
D	b		2733	- - - - - T C C T C T T G C G T A C C T - - - - -	2719
Q	y		335	A l a T h r T r p T h r A r g I l e T r p A s p T r p T h r S e r T y r P r o A s n A r g S e r L e u A r g T y r V a l	354
D	b		2718	- - - - - T T C G G T G C A T T T G G A T C A A A T T C A T C C T T A T G C C T	2683
Q	y		355	L e u A p I l e S e r A l a G l u P r o - - - - - T r p L e u T h r P h e	365
D	b		2682	G G C T T G A T T T C T T C G C C A C C A T A A T G A A C A G A T T T C A T C C A C T G G T T G T T T T G T T A C T A C T T	2623
Q	y		366	G l y V a l G l i n P r o A s n P r o V a l P r o S e r P r o L y s L e u G l y T r p M e t A s p G l u A l a M e t	385
D	b		2622	T C T G T T G G T T C A C C T T T C G C C A C T T T T T C C C C T G T T A A T G G G T T C T T A G T T G T T G G T G T T	2563
Q	y		386	A l a I l e - - - - - A s p P r o P h e A s n S e r A s p A r g M e t L e u T y r G l y T h r - - - - -	399
D	b		2562	G T A A T T G T T T T T C C T G T T C A C C T T T T T T T G A A C G A C T T T C T C T G T A C C T G C G C T A A G	2503
Q	y		400	- - - - - G l y A l a T h r L e u T y r A l a T h r A s n A s p L e u T h r L y s T r p A s p - - - - -	413
D	b		2502	T T T G G A T C A A A T T C G G G T T T T T A T C A A A C G G A A T T T C T C C G T T G A C G T A A C T G A T C	2443

Search completed: March 2, 2006, 13:23:14
Job time : 927.098 secs

```
Qy 414 ---SerGlyGlyGlnIleHisIleAlaProMetValIysGlyLeuGluGluThrAlaVal 432
Db 2442 CCATCAACTGGACCATATTTTGTACATCATCATCTGGTGAAGTAACCTACTTCGCCTGTA 2383
Qy 433 AsnAspLeuIleSerPro-----ProSerGlyAla--- 442
Db 2382 TCAGGATTTTAACTCTCTGGTTTACCTGGAACGCTCTTCGGTACCTTTTCGGTGCAATTT 2323
Qy 443 -----ProLeuIleSerAlaLeuGly 449
Db 2322 GGATCAAAATTCATCTTATGCGCTGGCTTGATTTCTTCGCCACCATATTTCTGTGATTCA 2263
Qy 450 AspLeuGlyGlyPheThrHisAlaAspValThrAlaValProSerThrIlePheThrSer 469
Db 2262 TCTACTGGTTGTTTGTACTTTTCTCTGTGTTCA--CCTTCGCCAACTTTT--TCC 2209
Qy 470 ProVal-----PheThrThrGlyThrSerValAspTyrAlaGluLeuAsnProSer 486
Db 2208 CCTGTTAATGGGTTCTTAGTTGTTGGTGTGTAATGTTTGTCTCTGGTTCACCTTTC 2149
Qy 487 IleIleValArgAla----- 491
Db 2148 TGTTTAACGCGCTCTTTACCTGGTTTAAATCAGATTGAATTCACGTTTCTTGTCTGAAT 2089
Qy 492 -----GlySerPheAsp-----ProSerSerGlnProAsnAspArgHisVal 505
Db 2088 GGAATTTCTTCGGTTGACGTGATCGGATCTCCATCAACTGGACCA----- 2044
Qy 506 AlaPheSerThrAspGlyGlyLysAsnTrpPheGlnGlySerGluProGlyGlyValThr 525
Db 2043 -----TATTTTGTCAATCATCCACCTGCTGGTGTGACT 2011
Qy 526 ThrGlyGlyThrValAlaAlaSerAlaAspGlySerArgPheValTrpAlaProGlyAsp 545
Db 2010 ACTTCGCGCTGTA-----TCAGGGTTTITA---ACTCCTGGTTTA 1975
Qy 546 ProGlyGlnProValValTyrAlaValGlyPheGlyAsnSerTrpAlaAlaSerGlnGly 565
Db 1974 CCTGGAACG-----TCCCTTGGCTACCTTTCGGTGCA 1942
Qy 566 ValProAlaAsnAlaGlnIleArgSerAspArgValAsnProLysThrPheTyrAlaLeu 585
Db 1941 TTTGGATCAAAATTCATCTTATGCGCTGGCTTGATTTCTTCGCCACCATATTTCTGTGATT 1882
Qy 586 SerAsn---GlyThrPheTyrArgSerThrAspGly-----GlyValThrPheGlnPro 602
Db 1881 TCATCTACTGTTGTTTGTATTTTCTGTGTTGGTTCACTTCGCCCAACTTTTTCGCCCT 1822
Qy 603 ValAlaAlaGlyLeuProSerSerGlyAlaValGlyValMetPheHisAlaValProGly 622
Db 1821 GTTAATGGGTTCTTA-----GTTGTTGGTGTGTAATTTGTTTGTCTCTGGT 1774
Qy 623 LysGluGlyAspLeuTrpLeuAlaAlaSerSerGlyLeuTyrHisSerThrAsnGlyGly 642
Db 1773 -----TCACCTTTTGTGTTAAACAGCTCTTCACCTGGTTTAAATCA-----GGATTG 1726
Qy 643 SerSerTrpSerAlaIleThrGlyValSerSerAlaValAsnValGlyPheGlyLysSer 662
Db 1725 AATTCACGTTTCTGTGTCGAATGGAATTTCTTCGTTGACGTGATCGGATCTCCATCAACT 1666
Qy 663 AlaProGlySerSerTyrProAlaValPheValValGlyThrIleGlyGlyValThrGly 682
Db 1665 GGACCA-----TATTTTGTCAATCATCCACAGGTGGAGTAACCTACT 1624
Qy 683 AlaTyrArgSerAsp-----AspCysGlyThrThrTrp 693
Db 1623 TCGCCTGTATCAGGATTTTAAACCCCGCGCTTACCTGGTTGCGTGTGTTGACTACCTTT 1564
Qy 694 ValLeuIleAsnAspAspGlnHisGlnTyrGly 704
Db 1563 CGGTGCATTTG-GATCAAAATTCATCCTTATGGC 1532
```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 13:55:49 ; Search time 5.12914 Seconds
(without alignments)
1434.576 Million cell updates/sec

Title: US-09-917-376-4

Perfect score: 469

Sequence: 1 VSGGVKQVQKNDAPGDNQ.....RASFGSVNPATPTADTYLQX 89

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA*

- 1: /cgn2_6/ptodata/1/iaa/5 COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/6 COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/H COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/RE COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	195.5	41.7	616	2	US-09-136-574A-47
2	194.5	41.5	1751	2	US-09-136-574A-44
3	191.5	40.8	1426	2	US-09-136-574A-43
4	172	36.7	700	1	US-07-862-588B-2
5	165.5	35.3	551	1	US-09-033-537A-1
6	165	35.2	167	4	PCT-US95-13813-9
7	165	35.2	476	2	US-09-339-159B-4
8	165	35.2	493	2	US-09-198-956-10
9	165	35.2	493	2	US-09-198-955A-12
10	165	35.2	493	2	US-09-694-531-12
11	165	35.2	493	2	US-09-670-141-10
12	165	35.2	493	2	US-10-072-152-12
13	155.5	33.2	1352	2	US-09-784-554B-2
14	150.5	32.1	1350	2	US-09-784-554B-4
15	120.5	25.7	531	1	US-07-862-588B-7
16	114	24.3	162	1	US-08-048-164A-2
17	114	24.3	162	1	US-08-460-462-2
18	114	24.3	162	1	US-08-460-457-2
19	114	24.3	162	1	US-08-460-458-2
20	114	24.3	162	1	US-08-460-455-2
21	114	24.3	162	1	US-08-330-394A-2
22	114	24.3	163	2	US-09-006-636-7
23	114	24.3	163	2	US-09-008-632-7
24	114	24.3	163	2	US-09-325-274-7
25	113	24.1	382	2	US-09-277-716-22
26	113	24.1	382	2	US-09-609-161B-22
27	112	23.9	154	1	US-08-330-394A-29

28	112	23.9	156	1	US-08-330-394A-22	Sequence 22, Appl
29	72.5	15.5	272	2	US-09-902-540-11056	Sequence 11056, A
30	68.5	14.6	500	2	US-10-104-047-2825	Sequence 2825, Ap
31	65.5	14.0	1566	2	US-09-581-472B-2	Sequence 2, Appli
32	65	13.9	474	2	US-09-248-796A-25524	Sequence 25524, A
33	64	13.6	143	2	US-09-301-593-26	Sequence 26, Appl
34	64	13.6	428	2	US-09-118-319-5	Sequence 5, Appli
35	64	13.6	464	1	US-08-353-400-36	Sequence 36, Appl
36	63.5	13.6	472	2	US-09-301-593-30	Sequence 30, Appl
37	63.5	13.5	1785	2	US-09-341-587-3	Sequence 3, Appli
38	63	13.4	453	2	US-09-301-593-18	Sequence 18, Appl
39	63	13.4	472	2	US-09-301-593-43	Sequence 43, Appl
40	63	13.4	718	2	US-09-328-352-4640	Sequence 4640, Ap
41	63	13.4	1133	2	US-09-902-540-12243	Sequence 12243, A
42	63	13.4	1581	2	US-09-110-517-2	Sequence 2, Appli
43	62.5	13.3	288	2	US-09-423-439-38	Sequence 38, Appl
44	62.5	13.3	389	2	US-09-902-540-14347	Sequence 14347, A
45	62.5	13.3	445	1	US-08-353-400-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1

US-09-136-574A-47

; Sequence 47, Application US/09136574A

; Patent No. 6294366

; GENERAL INFORMATION:

; APPLICANT: Farrington, Graham K.

; Anderson, Paige

; Gibbs, Moreland

; Bergquist, Peter

; Daniels, Roy

; Morgan, Hugh W.

; Williams, Diane P.

TITLE OF INVENTION: Compositions and Methods for Treating Cellulose Containing Fabrics Using Truncated Cellulase Enzyme Compositions

NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Center, P.O. Box 457

CITY: Spring House

STATE: PA

COUNTRY: USA

ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/136,574A

FILING DATE: 19-Aug-1998

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/932,571

FILING DATE: September 19, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Bak, Mary E.

REGISTRATION NUMBER: 31,215

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-540-9200

TELEFAX: 215-540-5818

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 47:

SEQUENCE CHARACTERISTICS:

LENGTH: 616 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 47:

US-07-862-588B-2
; Sequence 2, Application US/07862588B
; Patent No. 5916796
; GENERAL INFORMATION:
; APPLICANT: Joergensen, Per Linnaa
; APPLICANT: Sch lein, Martin
; APPLICANT: Hansen, Christian
; TITLE OF INVENTION: An Enzyme Exhibiting Cellulase Activity
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 59167960 No. 5916796disk of No. 5916796th America, Inc.
; STREET: 405 Lexington Avenue, 62nd floor
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/862,588B
; FILING DATE: 19920727
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 164/90
; FILING DATE: 19-JAN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/DK91/00013
; FILING DATE: 18-JAN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Zelson, Steve T. / Lambiris, Elias J.
; REGISTRATION NUMBER: 30,335 / 33,728
; REFERENCE/DOCKET NUMBER: 3425.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 700 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-862-588B-2

Query Match 36.7%; Score 172; DB 1; Length 700;
Best Local Similarity 44.9%; Pred. No. 1.5e-10;
Matches 40; Conservative 18; Mismatches 27; Indels 4; Gaps 3;
QY 1 VSGGVKVOYKND-SAPGDNOIKPGLQLVNTGSSVDLSTVTYVYWFTRDGGSSTLVYNC 59
DB 549 VNSDLVVOYKGDNRNATDNOIKPHFNQNGTSPVDLSLTLRYFTKD-SSAAMNGWI 607
QY 60 DWAAMGCGNIRASFGSVNPATPTADTYLQ 88
DB 608 DWAKLGGNIOISFGNHA--DSDTYAE 634

RESULT 5
US-09-033-537A-1
; Sequence 1, Application US/09033537A
; Patent No. 5958083
; GENERAL INFORMATION:
; APPLICANT: Onishi, Masahiro
; APPLICANT: Fich, Merete
; APPLICANT: Toft, Annette Hanne
; APPLICANT: Sh lein, Martin
; TITLE OF INVENTION: Prevention Of Back-Staining
; TITLE OF INVENTION: In Stone Washing
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5958083o No. 5958083disk of No. 5958083th America, Inc.

US-07-862-588B-2
; Sequence 2, Application US/07862588B
; Patent No. 5916796
; GENERAL INFORMATION:
; APPLICANT: Joergensen, Per Linnaa
; APPLICANT: Sch lein, Martin
; APPLICANT: Hansen, Christian
; TITLE OF INVENTION: An Enzyme Exhibiting Cellulase Activity
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 59167960 No. 5916796disk of No. 5916796th America, Inc.
; STREET: 405 Lexington Avenue, 62nd floor
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/862,588B
; FILING DATE: 19920727
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 164/90
; FILING DATE: 19-JAN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/DK91/00013
; FILING DATE: 18-JAN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Zelson, Steve T. / Lambiris, Elias J.
; REGISTRATION NUMBER: 30,335 / 33,728
; REFERENCE/DOCKET NUMBER: 3425.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 700 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-862-588B-2

Query Match 36.7%; Score 172; DB 1; Length 700;
Best Local Similarity 44.9%; Pred. No. 1.5e-10;
Matches 40; Conservative 18; Mismatches 27; Indels 4; Gaps 3;
QY 1 VSGGVKVOYKND-SAPGDNOIKPGLQLVNTGSSVDLSTVTYVYWFTRDGGSSTLVYNC 59
DB 549 VNSDLVVOYKGDNRNATDNOIKPHFNQNGTSPVDLSLTLRYFTKD-SSAAMNGWI 607
QY 60 DWAAMGCGNIRASFGSVNPATPTADTYLQ 88
DB 608 DWAKLGGNIOISFGNHA--DSDTYAE 634

RESULT 5
US-09-033-537A-1
; Sequence 1, Application US/09033537A
; Patent No. 5958083
; GENERAL INFORMATION:
; APPLICANT: Onishi, Masahiro
; APPLICANT: Fich, Merete
; APPLICANT: Toft, Annette Hanne
; APPLICANT: Sh lein, Martin
; TITLE OF INVENTION: Prevention Of Back-Staining
; TITLE OF INVENTION: In Stone Washing
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5958083o No. 5958083disk of No. 5958083th America, Inc.

STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: U.S.A.
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/033,537A
FILING DATE: 02-MAR-1998
CLASSIFICATION: 008
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 0993/95
FILING DATE: 08-SEP-1995
APPLICATION NUMBER: PCT/DK96/00364
FILING DATE: 03-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4492.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 551 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-033-537A-1

Query Match 35.3%; Score 165.5; DB 1; Length 551;
Best Local Similarity 37.9%; Pred. No. 6e-10;
Matches 33; Conservative 22; Mismatches 29; Indels 3; Gaps 2;
QY 2 SGGVKVOYKND-SAPGDNOIKPGLQLVNTGSSVDLSTVTYVYWFTRDGGSSTLVYNC 61
DB 402 TGNLVVOYKVGDTSDTDNQMKPSFNKNGTTPVNLGKLYRYFTKD-GTADMSASFDW 460
QY 62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
DB 461 AQIGASNVSAF--ANFTGSNTDTTVE 485

RESULT 6
PCT-US95-13813-9
; Sequence 9, Application PC/TUS9513813
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co. Ltd.
; APPLICANT: Ramot University Authority for Applied
; APPLICANT: Research and Industrial Development Ltd.
; APPLICANT: Technion Research and Development Foundation Ltd.
; APPLICANT: Bayer, Edward A.
; APPLICANT: Morag, Ely
; APPLICANT: Wilchek, Meir
; APPLICANT: Lamed, Raphael
; APPLICANT: Shoham, Yuval
; TITLE OF INVENTION: MODIFIED CELLULOSE-BINDING DOMAIN (CBD)
; TITLE OF INVENTION: PROTEINS AND USE THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brody and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13813
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: BAYER=3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 167 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-13813-9

Query Match 35.2%; Score 165; DB 4; Length 167;
Best Local Similarity 36.1%; Pred. No. 1.6e-10;
Matches 35; Conservative 23; Mismatches 29; Indels 10; Gaps 2;

Qy 1 VSGGVKVOYKNDSAPGDNQIKPGLQLVNTGSSVDLSVTYVYWFTRDGGSSTLVYVNC 60
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
5 VSGNLKVEFYNSPDTTNSINPQKVTNTGSSAIDLSKLTLYRYVTVDGQKQTFW-CD 63
Qy 61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
64 HAALIGNSYNGITSNVKGTFVKMSSTNNADTYLE 100

RESULT 7
US-09-339-159B-4
; Sequence 4, Application US/09339159B
; Patent No. 6566114
; GENERAL INFORMATION:
; APPLICANT: Kauppinen, Markus
; APPLICANT: Schulein, Martin
; APPLICANT: Schnorr, Kirk
; APPLICANT: Andersen, Lene
; APPLICANT: Bjornvad, Mads
; TITLE OF INVENTION: No. 6566114el Mannanases
; FILE REFERENCE: 5440.204-US
; CURRENT APPLICATION NUMBER: US/09/339,159B
; CURRENT FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Bacillus
US-09-339-159B-4

Query Match 35.2%; Score 165; DB 2; Length 476;
Best Local Similarity 36.1%; Pred. No. 5.7e-10;
Matches 35; Conservative 23; Mismatches 29; Indels 10; Gaps 2;

Qy 1 VSGGVKVOYKNDSAPGDNQIKPGLQLVNTGSSVDLSVTYVYWFTRDGGSSTLVYVNC 60
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
314 VSGNLKVEFYNSPDTTNSINPQKVTNTGSSAIDLSKLTLYRYVTVDGQKQTFW-CD 372
Qy 61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
373 HAALIGNSYNGITSNVKGTFVKMSSTNNADTYLE 409

RESULT 8
US-09-198-956-10
; Sequence 10, Application US/09198956
; Patent No. 6165769
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schulein, Martin
; APPLICANT: Lange, Niels Erik K.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Schnorr, Kirk
; TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus
; FILE REFERENCE: 5377.200-US
; CURRENT APPLICATION NUMBER: US/09/198,956
; CURRENT FILING DATE: 1998-11-24
; EARLIER APPLICATION NUMBER: 1344/97
; EARLIER FILING DATE: 1997-11-24
; EARLIER APPLICATION NUMBER: 60/067,240
; EARLIER FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-09-198-956-10

Query Match 35.2%; Score 165; DB 2; Length 493;
Best Local Similarity 36.1%; Pred. No. 5.9e-10;
Matches 35; Conservative 23; Mismatches 29; Indels 10; Gaps 2;

Qy 1 VSGGVKVOYKNDSAPGDNQIKPGLQLVNTGSSVDLSVTYVYWFTRDGGSSTLVYVNC 60
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
331 VSGNLKVEFYNSPDTTNSINPQKVTNTGSSAIDLSKLTLYRYVTVDGQKQTFW-CD 389
Qy 61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
390 HAALIGNSYNGITSNVKGTFVKMSSTNNADTYLE 426

RESULT 9
US-09-198-955A-12
; Sequence 12, Application US/09198955A
; Patent No. 6187580
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schulein, Martin
; APPLICANT: Lange, Niels E.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Moller, Soren
; APPLICANT: Glad, Sanne O. S.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Schnorr, Kirk
; APPLICANT: Kongsbak, Lars
; TITLE OF INVENTION: No. 6187580el Pectate Lyases
; FILE REFERENCE: 5378.200-US
; CURRENT APPLICATION NUMBER: US/09/198,955A
; CURRENT FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: 1343/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 1344/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/067,249
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 60/067,240
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 09/073,684
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 09/184,217
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 493
```

```
; TYPE: PRT
; ORGANISM: Clostridium thermocellum
US-09-198-955A-12

Query Match      35.2%; Score 165; DB 2; Length 493;
Best Local Similarity 36.1%; Pred. No. 5.9e-10;
Matches 35; Conservative 23; Mismatches 29; Indels 10; Gaps 2;

QY 1 VSGGVKQYKNDGAPGNQIKPGLQLVNTGSSVDLSTVTVRYWFTRDGSGSTLVYVNC 60
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 331 VSGNLKVEFYNSPDDTTNSINPQKVTNTGSSAIDLKLTLYYYTVDGKQDTFW-CD 389
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 390 HAAIIGNSYNGITSNVKGTFFVKMSSSTNNADTYLE 426
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 10
; Sequence 12, Application US/09694531
; Patent No. 6368843
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schulein, Martin
; APPLICANT: Lange, Niels E.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Moller, Soren
; APPLICANT: Glad, Sanne O. S.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Schnorr, Kirk
; APPLICANT: Kongsbak, Lars
; TITLE OF INVENTION: No. 6368843el Pectate Lyases
; FILE REFERENCE: 5378.200-US
; CURRENT APPLICATION NUMBER: US/09/694,531
; CURRENT FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 09/198,955
; PRIOR FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: 1343/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 1344/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/067,249
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 60/067,240
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 09/073,684
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 09/184,217
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Clostridium thermocellum
US-09-694-531-12

Query Match      35.2%; Score 165; DB 2; Length 493;
Best Local Similarity 36.1%; Pred. No. 5.9e-10;
Matches 35; Conservative 23; Mismatches 29; Indels 10; Gaps 2;

QY 1 VSGGVKQYKNDGAPGNQIKPGLQLVNTGSSVDLSTVTVRYWFTRDGSGSTLVYVNC 60
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 331 VSGNLKVEFYNSPDDTTNSINPQKVTNTGSSAIDLKLTLYYYTVDGKQDTFW-CD 389
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 390 HAAIIGNSYNGITSNVKGTFFVKMSSSTNNADTYLE 426
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 11
US-09-670-141-10
; Sequence 10, Application US/09670141
```

```
; Patent No. 6429000
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schulein, Martin
; APPLICANT: Lange, Niels E.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Schnorr, Kirk
; TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus
; FILE REFERENCE: 5377.200-US
; CURRENT APPLICATION NUMBER: US/09/670,141
; CURRENT FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: 09/198,956
; PRIOR FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: 1344/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/067,240
; PRIOR FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-09-670-141-10

Query Match      35.2%; Score 165; DB 2; Length 493;
Best Local Similarity 36.1%; Pred. No. 5.9e-10;
Matches 35; Conservative 23; Mismatches 29; Indels 10; Gaps 2;

QY 1 VSGGVKQYKNDGAPGNQIKPGLQLVNTGSSVDLSTVTVRYWFTRDGSGSTLVYVNC 60
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 331 VSGNLKVEFYNSPDDTTNSINPQKVTNTGSSAIDLKLTLYYYTVDGKQDTFW-CD 389
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 61 WAAM-----GCGNIRASFGSVNPATPTADTYLQ 88
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 390 HAAIIGNSYNGITSNVKGTFFVKMSSSTNNADTYLE 426
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 12
US-10-072-152-12
; Sequence 12, Application US/10072152
; Patent No. 6677147
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schulein, Martin
; APPLICANT: Lange, Niels E.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Moller, Soren
; APPLICANT: Glad, Sanne O. S.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Schnorr, Kirk
; APPLICANT: Kongsbak, Lars
; TITLE OF INVENTION: No. 6677147el Pectate Lyases
; FILE REFERENCE: 5378.200-US
; CURRENT APPLICATION NUMBER: US/10/072,152
; CURRENT FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: US/09/198,955
; PRIOR FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: 1343/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 1344/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/067,249
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 60/067,240
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 09/073,684
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 09/184,217
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
```

```
; SEQ ID NO 12
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Clostridium thermocellum
US-10-072-152-12

Query Match      35.2%; Score 165; DB 2; Length 493;
Best Local Similarity 36.1%; Pred. No. 5.9e-10;
Matches 35; Conservative 23; Mismatches 29; Indels 10; Gaps 2;

Qy 1 VSGGVKQYKNDSPAGDNOIKPGLQVLVNTGSSVDLSLTVTVRYWFTRDGSSSTLVNCD 60
Db 331 VSGNLKVEFYNSPDSITNSINPDKVNTNTSSAIDLSKULFRIYYTVGQKQOTFW-CD 389

Qy 61 WAAM-----CCGNIRASFGSVNPATPTADTYLQ 88
Db 390 HAAIIGSGSYNGITSNVKGFVKMSSSTNNADIYLE 426

RESULT 13
US-09-784-554B-2
; Sequence 2, Application US/09784554B
; Patent No. 6815192
; GENERAL INFORMATION:
; APPLICANT: Schnorr, Kirk
; APPLICANT: Jorgensen, Per Lina
; APPLICANT: Schulein, Martin
; TITLE OF INVENTION: FAMILY 44 XYLOGLUCANASES
; FILE REFERENCE: 10017.200-US
; CURRENT APPLICATION NUMBER: US/09/784.554B
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1352
; TYPE: PRT
; ORGANISM: Paenibacillus polymyxa
US-09-784-554B-2

Query Match      33.2%; Score 155.5; DB 2; Length 1352;
Best Local Similarity 31.8%; Pred. No. 2.3e-08;
Matches 27; Conservative 21; Mismatches 36; Indels 1; Gaps 1;

Qy 4 GVKVQYKNDSPAGDNOIKPGLQVLVNTGSSVDLSLTVTVRYWFTRDGSSSTLVNCDWAA 63
Db 1200 GLLQYRTADTKYNDHNLNPFQVLVNGTTSIPINELKIRYYITDGDREQ-TFNCDYAT 1258

Qy 64 MCGGNIRASFGSVNPATPTADTYLQ 88
Db 1259 LSCSKLNGKLVKMDKAATGADYILE 1283

RESULT 14
US-09-784-554B-4
; Sequence 4, Application US/09784554B
; Patent No. 6815192
; GENERAL INFORMATION:
; APPLICANT: Schnorr, Kirk
; APPLICANT: Jorgensen, Per Lina
; APPLICANT: Schulein, Martin
; TITLE OF INVENTION: FAMILY 44 XYLOGLUCANASES
; FILE REFERENCE: 10017.200-US
; CURRENT APPLICATION NUMBER: US/09/784.554B
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1350
; TYPE: PRT
; ORGANISM: Paenibacillus polymyxa
US-09-784-554B-4

Query Match      32.1%; Score 150.5; DB 2; Length 1350;
```

```
Best Local Similarity 31.8%; Pred. No. 8.1e-08;
Matches 27; Conservative 21; Mismatches 36; Indels 1; Gaps 1;

Qy 4 GVKVQYKNDSPAGDNOIKPGLQVLVNTGSSVDLSLTVTVRYWFTRDGSSSTLVNCDWAA 63
Db 1199 GLVQYRTADTKYNDHNLNPFQVLVNGTTSIPINELKIRYYITDGDREQ-TFNCDYAV 1257

Qy 64 MCGGNIRASFGSVNPATPTADTYLQ 88
Db 1258 LSCSKLNGKLVKMDKAATGADYILE 1282

RESULT 15
US-07-862-588B-7
; Sequence 7, Application US/07862588B
; Patent No. 5916796
; GENERAL INFORMATION:
; APPLICANT: Joergensen, Per Linaa
; APPLICANT: Sch lein, Martin
; APPLICANT: Hansen, Christian
; TITLE OF INVENTION: An Enzyme Exhibiting Cellulase Activity
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5916796o No. 5916796disk of No. 5916796th America, Inc.
; STREET: 405 Lexington Avenue, 62nd floor
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/862.588B
; FILING DATE: 19920727
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 164/90
; FILING DATE: 19-JAN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/DK91/00013
; FILING DATE: 18-JAN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Zelson, Steve T. / Lambiris, Elias J.
; REGISTRATION NUMBER: 30,335 / 33,728
; REFERENCE/DOCKET NUMBER: 3425.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 531 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-862-588B-7

Query Match      25.7%; Score 120.5; DB 1; Length 531;
Best Local Similarity 32.9%; Pred. No. 5.2e-05;
Matches 28; Conservative 18; Mismatches 36; Indels 3; Gaps 2;

Qy 2 SGGVKQYKNDSPAGDNOIKPGLQVLVNTGSSVDLSLTVTVRYWFTRDGSSSTLVNCDW 61
Db 432 TGNLVQYKVGDTSATDNQMKPSFNKNGTTPVNLGLKXXXXKD-GPADMSCSIDW 490

Qy 62 AAMCGGNIRASFGSVNPATPTADTY 86
Db 491 AQIGRTNVLAF--ANFTGSNTDTY 513

Search completed: March 2, 2006, 13:58:00
```

Job time : 6.12914 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 14:20:38 ; Search time 17.1446 Seconds
(without alignments)
2169.009 Million cell updates/sec

Title: US-09-917-376-4

Perfect score: 469

Sequence: 1 VSGGVKQVKNDSAPGDNQ.....RASFGSVNPATPTADTYLQX 89

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	469	100.0	88	3	US-09-917-376-5
2	469	100.0	88	4	US-10-155-400-5
3	469	100.0	89	3	US-09-917-376-4
4	469	100.0	89	4	US-10-155-400-4
5	469	100.0	154	3	US-09-917-378-4
6	469	100.0	762	3	US-09-917-378-1
7	466	99.4	150	3	US-09-917-384-5
8	466	99.4	150	3	US-09-917-383-5
9	466	99.4	1043	3	US-09-917-384-6
10	466	99.4	1043	3	US-09-917-383-6
11	466	99.4	1228	3	US-09-917-384-1
12	466	99.4	1228	3	US-09-917-383-1
13	462	98.5	957	3	US-09-917-376-1
14	462	98.5	957	4	US-10-155-400-1
15	257	54.8	741	4	US-10-156-761-8100
16	190	40.5	508	4	US-10-369-493-23151
17	188	40.1	477	4	US-10-466-208-12
18	188	40.1	496	4	US-10-466-208-8
19	188	40.1	677	4	US-10-433-577-35
20	180	38.4	1621	4	US-10-185-990-10
21	165	35.2	167	5	US-10-933-404-4
22	165	35.2	476	4	US-10-372-054-4
23	165	35.2	493	4	US-10-072-152-12
24	165	35.2	493	4	US-10-655-433-12
25	165	35.2	599	3	US-09-955-555A-29
26	155.5	33.2	1352	3	US-09-784-554B-2
27	155.5	33.2	1352	5	US-10-896-555-2

28	150.5	32.1	1350	3	US-09-784-554B-4	Sequence 4, Appli
29	150.5	32.1	1350	5	US-10-896-555-4	Sequence 4, Appli
30	130.5	27.8	1483	4	US-10-282-122A-51483	Sequence 51483, A
31	114	24.3	256	4	US-10-261-446-6	Sequence 6, Appli
32	114	24.3	256	4	US-10-261-445B-6	Sequence 6, Appli
33	114	24.3	256	5	US-10-782-234-6	Sequence 6, Appli
34	113	24.1	382	3	US-09-808-898-22	Sequence 22, Appli
35	105.5	22.5	163	4	US-10-460-524-2	Sequence 2, Appli
36	78.5	16.7	1049	4	US-10-282-122A-49900	Sequence 49900, A
37	75.5	16.1	346	4	US-10-437-963-179234	Sequence 179234,
38	75.5	16.1	618	4	US-10-211-462-223	Sequence 223, App
39	75.5	16.1	618	5	US-10-723-860-4039	Sequence 4039, Ap
40	75.5	16.1	618	5	US-10-756-149-5713	Sequence 5713, Ap
41	75.5	16.1	986	3	US-09-747-835A-53	Sequence 53, Appl
42	75.5	16.1	986	4	US-10-120-604-101	Sequence 101, App
43	75.5	16.1	986	4	US-10-225-567A-406	Sequence 406, App
44	75.5	16.1	986	4	US-10-398-458-3	Sequence 3, Appli
45	75.5	16.1	986	4	US-10-312-312-53	Sequence 53, Appl

ALIGNMENTS

RESULT 1

US-09-917-376-5
; Sequence 5, Application US/09917376
; Publication No. US20040038334A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOLYTICUS
; FILE REFERENCE: 40197.4US01
; CURRENT APPLICATION NUMBER: US/09/917,376
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
; FEATURE:
; OTHER INFORMATION: Carbohydrate binding domain
US-09-917-376-5

Query Match		100.0%;	Score 469;	DB 3;	Length 88;
Best Local Similarity		100.0%;	Pred. No. 3.6e-48;		
Matches		88;	Conservative	0;	Mismatches
				0;	Indels
				0;	Gaps
				0;	
QY	1	VSGGVKQVKNDSAPGDNQIKPGLQVNTGSSVDLSTVTYVYFTRDGGSTLVNCD	60		
Db	1	VSGGVKQVKNDSAPGDNQIKPGLQVNTGSSVDLSTVTYVYFTRDGGSTLVNCD	60		
QY	61	WAAMGCGNIRASFGSVNPATPTADTYLQ	88		
Db	61	WAAMGCGNIRASFGSVNPATPTADTYLQ	88		

RESULT 2

US-10-155-400-5
; Sequence 5, Application US/10155400
; Publication No. US2003010898A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOLYTICUS
; FILE REFERENCE: NREL 01-36A
; CURRENT APPLICATION NUMBER: US/10/155,400


```
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
; FEATURE:
; OTHER INFORMATION: Carbohydrate binding domain
US-10-155-400-5

Query Match          100.0%; Score 469; DB 4; Length 88;
Best Local Similarity 100.0%; Pred. No. 3.7e-48;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSGGVKVOYKNDSPAGDNQIKPGLQLVNTGSSVDLSTVTYRYWFTRDGGSSTLVNCD 60
Db 1 VSGGVKVOYKNDSPAGDNQIKPGLQLVNTGSSVDLSTVTYRYWFTRDGGSSTLVNCD 60

Qy 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
Db 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88

RESULT 3
US-09-917-376-4
; Sequence 4, Application US/09917376
; Publication No. US20040038334A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
; FILE REFERENCE: 40197.4US01
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
; NAME/KEY: MOD RES
; LOCATION: (89)
; OTHER INFORMATION: Any amino acid
; OTHER INFORMATION: Carbohydrate binding domain
US-09-917-376-4

Query Match          100.0%; Score 469; DB 3; Length 89;
Best Local Similarity 100.0%; Pred. No. 3.7e-48;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSGGVKVOYKNDSPAGDNQIKPGLQLVNTGSSVDLSTVTYRYWFTRDGGSSTLVNCD 60
Db 1 VSGGVKVOYKNDSPAGDNQIKPGLQLVNTGSSVDLSTVTYRYWFTRDGGSSTLVNCD 60

Qy 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
Db 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88

RESULT 4
US-10-155-400-4
; Sequence 4, Application US/10155400
; Publication No. US20030108988A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
```

```
; TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOLYTICUS
; FILE REFERENCE: NREL 01-36A
; CURRENT APPLICATION NUMBER: US/10/155,400
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (89)
; OTHER INFORMATION: Any amino acid
; FEATURE:
; OTHER INFORMATION: Carbohydrate binding domain
US-10-155-400-4

Query Match          100.0%; Score 469; DB 4; Length 89;
Best Local Similarity 100.0%; Pred. No. 3.7e-48;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSGGVKVOYKNDSPAGDNQIKPGLQLVNTGSSVDLSTVTYRYWFTRDGGSSTLVNCD 60
Db 1 VSGGVKVOYKNDSPAGDNQIKPGLQLVNTGSSVDLSTVTYRYWFTRDGGSSTLVNCD 60

Qy 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
Db 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88

RESULT 5
US-09-917-378-4
; Sequence 4, Application US/09917378
; Publication No. US20030119093A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT MANNANASE FROM ACIDOTHERMUS
; FILE REFERENCE: 40197.7US01
; CURRENT APPLICATION NUMBER: US/09/917,378
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Carbohydrate
; OTHER INFORMATION: binding domain
US-09-917-378-4

Query Match          100.0%; Score 469; DB 3; Length 154;
Best Local Similarity 100.0%; Pred. No. 7.1e-48;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSGGVKVOYKNDSPAGDNQIKPGLQLVNTGSSVDLSTVTYRYWFTRDGGSSTLVNCD 60
Db 1 VSGGVKVOYKNDSPAGDNQIKPGLQLVNTGSSVDLSTVTYRYWFTRDGGSSTLVNCD 60

Qy 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
Db 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88

RESULT 6
US-09-917-378-1
```

```
; Sequence 1, Application US/09917378
; Publication No. US20030119093A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT MANNANASE FROM ACIDOTHERMUS
; FILE REFERENCE: 40197.7US01
; CURRENT APPLICATION NUMBER: US/09/917,378
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 762
; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
US-09-917-378-1

Query Match      100.0%; Score 469; DB 3; Length 762;
Best Local Similarity 100.0%; Pred. No. 4.9e-47;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSGGVKQYKND SAPGNQIKPGLQLVNTGSSVDLSTVTTRYWFTDRDGSSTLVNCD 60
Db 455 VSGGVKQYKND SAPGNQIKPGLQLVNTGSSVDLSTVTTRYWFTDRDGSSTLVNCD 514

Qy 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
Db 515 WAAMCGNIRASFGSVNPATPTADTYLQ 542

RESULT 7
US-09-917-384-5
; Sequence 5, Application US/09917384
; Publication No. US20030096342A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE FROM ACIDOTHERMUS
; FILE REFERENCE: 40170.6US01
; CURRENT APPLICATION NUMBER: US/09/917,384
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Segment of
; OTHER INFORMATION: Guxa
US-09-917-384-5

Query Match      99.4%; Score 466; DB 3; Length 150;
Best Local Similarity 98.9%; Pred. No. 1.6e-47;
Matches 87; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSGGVKQYKND SAPGNQIKPGLQLVNTGSSVDLSTVTTRYWFTDRDGSSTLVNCD 60
Db 455 VSGGVKQYKND SAPGNQIKPGLQLVNTGSSVDLSTVTTRYWFTDRDGSSTLVNCD 514

Qy 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
Db 515 WAAMCGNIRASFGSVNPATPTADTYLQ 542

RESULT 7
US-09-917-384-5
; Sequence 5, Application US/09917384
; Publication No. US20030096342A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE FROM ACIDOTHERMUS
; FILE REFERENCE: 40170.6US01
; CURRENT APPLICATION NUMBER: US/09/917,384
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Segment of
; OTHER INFORMATION: Guxa
US-09-917-384-5

Query Match      99.4%; Score 466; DB 3; Length 150;
Best Local Similarity 98.9%; Pred. No. 1.6e-47;
Matches 87; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSGGVKQYKND SAPGNQIKPGLQLVNTGSSVDLSTVTTRYWFTDRDGSSTLVNCD 60
Db 455 VSGGVKQYKND SAPGNQIKPGLQLVNTGSSVDLSTVTTRYWFTDRDGSSTLVNCD 514

Qy 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
Db 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
```

```
RESULT 8
US-09-917-383-5
; Sequence 5, Application US/09917383
; Publication No. US20030104522A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE FROM ACIDOTHERMUS
; FILE REFERENCE: 40170.6US01
; CURRENT APPLICATION NUMBER: US/09/917,383
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Segment of
; OTHER INFORMATION: Guxa
US-09-917-383-5

Query Match      99.4%; Score 466; DB 3; Length 150;
Best Local Similarity 98.9%; Pred. No. 1.6e-47;
Matches 87; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSGGVKQYKND SAPGNQIKPGLQLVNTGSSVDLSTVTTRYWFTDRDGSSTLVNCD 60
Db 455 VSGGVKQYKND SAPGNQIKPGLQLVNTGSSVDLSTVTTRYWFTDRDGSSTLVNCD 514

Qy 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
Db 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88

RESULT 9
US-09-917-384-6
; Sequence 6, Application US/09917384
; Publication No. US20030096342A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE FROM ACIDOTHERMUS
; FILE REFERENCE: 40170.6US01
; CURRENT APPLICATION NUMBER: US/09/917,384
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Segment of
; OTHER INFORMATION: Guxa
US-09-917-384-6

Query Match      99.4%; Score 466; DB 3; Length 1043;
Best Local Similarity 98.9%; Pred. No. 1.7e-46;
Matches 87; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSGGVKQYKND SAPGNQIKPGLQLVNTGSSVDLSTVTTRYWFTDRDGSSTLVNCD 60
Db 477 VSGGVKQYKND SAPGNQIKPGLQLVNTGSSVDLSTVTTRYWFTDRDGSSTLVNCD 536
```

Qy 61 WAAMCGNIRASFGSNPATPTADTYLQ 88
Db 537 WAAMCGNIRASFGSNPATPTADTYLQ 564

RESULT 10
US-09-917-383-6
; Sequence 6, Application US/09917383
; Publication No. US20030104522A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE FROM ACIDOTHERMUS
; FILE REFERENCE: 40170.6US01
; CURRENT APPLICATION NUMBER: US/09/917,383
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Segment of Guxa
US-09-917-383-6

Query Match 99.4%; Score 466; DB 3; Length 1043;
Best Local Similarity 98.9%; Pred. No. 1.7e-46;
Matches 87; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSGGVKQYKNDSPAGDNQIKPGLQLVNTGSSVDLSTVTYRYWFTRDGSSSTLVNCD 60
Db 477 VSGGLKVQYKNDSPAGDNQIKPGLQLVNTGSSVDLSTVTYRYWFTRDGSSSTLVNCD 536

Qy 61 WAAMCGNIRASFGSNPATPTADTYLQ 88
Db 537 WAAMCGNIRASFGSNPATPTADTYLQ 564

RESULT 11
US-09-917-384-1
; Sequence 1, Application US/09917384
; Publication No. US20030096342A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE FROM ACIDOTHERMUS
; FILE REFERENCE: 40170.6US01
; CURRENT APPLICATION NUMBER: US/09/917,384
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1228
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Segment of Guxa
US-09-917-384-1

Query Match 99.4%; Score 466; DB 3; Length 1228;
Best Local Similarity 98.9%; Pred. No. 2e-46;
Matches 87; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSGGVKQYKNDSPAGDNQIKPGLQLVNTGSSVDLSTVTYRYWFTRDGSSSTLVNCD 60
Db 584 VSGGLKVQYKNDSPAGDNQIKPGLQLVNTGSSVDLSTVTYRYWFTRDGSSSTLVNCD 643

Qy 61 WAAMCGNIRASFGSNPATPTADTYLQ 88
Db 644 WAAMCGNIRASFGSNPATPTADTYLQ 671

RESULT 12
US-09-917-383-1
; Sequence 1, Application US/09917383
; Publication No. US20030104522A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE FROM ACIDOTHERMUS
; FILE REFERENCE: 40170.6US01
; CURRENT APPLICATION NUMBER: US/09/917,383
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1228
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Segment of Guxa
US-09-917-383-1

Query Match 99.4%; Score 466; DB 3; Length 1228;
Best Local Similarity 98.9%; Pred. No. 2e-46;
Matches 87; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSGGVKQYKNDSPAGDNQIKPGLQLVNTGSSVDLSTVTYRYWFTRDGSSSTLVNCD 60
Db 584 VSGGLKVQYKNDSPAGDNQIKPGLQLVNTGSSVDLSTVTYRYWFTRDGSSSTLVNCD 643

Qy 61 WAAMCGNIRASFGSNPATPTADTYLQ 88
Db 644 WAAMCGNIRASFGSNPATPTADTYLQ 671

RESULT 13
US-09-917-376-1
; Sequence 1, Application US/09917376
; Publication No. US20040038334A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
; FILE REFERENCE: 40197.4US01
; CURRENT APPLICATION NUMBER: US/09/917,376
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 957
; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (957)
; OTHER INFORMATION: Any amino acid

US-09-917-376-1

Query Match 98.5%; Score 462; DB 3; Length 957;
Best Local Similarity 97.7%; Pred. No. 4.5e-46;
Matches 86; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 VSGGVKVOYKNDSPAGDNQIKPGLQVNTGSSVDLSVTVTYVWFTTRDGGSSSTLVYVNC 60
Db 869 VSGGVKVOYKNDSPAGDNQIKPGLQVNTGSSVDLSVTVTYVWFTTRDGGSSSTLVYVNC 928
QY 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
Db 929 WAAIGCGNIRASFGSVNPATPTADTYLQ 956

RESULT 14

US-10-155-400-1-
; Sequence 1, Application US/10155400
; Publication No. US20030108988A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOGLYTICUS
; FILE REFERENCE: NREL 01-36A
; CURRENT APPLICATION NUMBER: US/10/155,400
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 957
; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (957)
; OTHER INFORMATION: Any amino acid
US-10-155-400-1

Query Match 98.5%; Score 462; DB 4; Length 957;
Best Local Similarity 97.7%; Pred. No. 4.5e-46;
Matches 86; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 VSGGVKVOYKNDSPAGDNQIKPGLQVNTGSSVDLSVTVTYVWFTTRDGGSSSTLVYVNC 60
Db 869 VSGGVKVOYKNDSPAGDNQIKPGLQVNTGSSVDLSVTVTYVWFTTRDGGSSSTLVYVNC 928
QY 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
Db 929 WAAIGCGNIRASFGSVNPATPTADTYLQ 956

RESULT 15

US-10-156-761-8100
; Sequence 8100, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697

; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8100
; LENGTH: 741
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8100
Query Match 54.8%; Score 257; DB 4; Length 741;
Best Local Similarity 56.3%; Pred. No. 1.2e-21;
Matches 49; Conservative 13; Mismatches 25; Indels 0; Gaps 0;
QY 2 SGGVKVOYKNDSPAGDNQIKPGLQVNTGSSVDLSVTVTYVWFTTRDGGSSSTLVYVNC 61
Db 591 SGGKVLVYKNDSSATDNAIRPGLRIYNTGSGSLDLSKVTARYYFTRDGGSGPTVNAWCDY 650
QY 62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
Db 651 AAVGCSNVSLKVVPLTTPVPGADAYLE 677
Search completed: March 2, 2006, 14:26:55
Job time: 18.1446 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 14:27:20 ; Search time 1.8047 Seconds
(without alignments)
986.322 Million cell updates/sec

Title: US-09-917-376-4

Perfect score: 469

Sequence: 1 VSGGVKQYKNDSPAGDQ.....RASFGSVNPATPTADTYLQX 89

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 135339 seqs, 20000136 residues

Total number of hits satisfying chosen parameters: 135339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:
1: /cgn2_6/ptodata/2/pubaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/2/pubaa/US06_NEW_PUB pep.*
3: /cgn2_6/ptodata/2/pubaa/US07_NEW_PUB pep.*
4: /cgn2_6/ptodata/2/pubaa/PCT_NEW_PUB pep.*
5: /cgn2_6/ptodata/2/pubaa/US09_NEW_PUB pep.*
6: /cgn2_6/ptodata/2/pubaa/US10_NEW_PUB pep.*
7: /cgn2_6/ptodata/2/pubaa/US11_NEW_PUB pep.*
8: /cgn2_6/ptodata/2/pubaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	163	34.8	499	6	US-10-517-939-328
2	157	33.5	534	6	US-10-510-386-230
3	113	24.1	382	7	US-11-179-411-22
4	113	24.1	382	7	US-11-175-766-22
5	108	23.0	738	6	US-10-517-939-344
6	72	15.4	210	7	US-11-170-653-51
7	70	14.9	5291	7	US-11-052-554A-281
8	68.5	14.6	500	7	US-11-072-512-2825
9	68	14.5	755	6	US-10-517-939-330
10	64.5	13.8	389	7	US-11-207-626A-26
11	62	13.2	389	7	US-11-207-626A-20
12	61	13.0	350	6	US-10-517-939-14
13	61	13.0	1889	7	US-11-102-476-46
14	60.5	12.9	806	7	US-11-098-686-10278
15	60.5	12.9	983	7	US-11-087-099-7483
16	60	12.8	86	7	US-11-102-476-24
17	60	12.8	91	7	US-11-102-476-9
18	60	12.8	919	7	US-11-074-176-284
19	60	12.8	1225	7	US-11-102-476-2
20	60	12.8	1424	7	US-11-102-476-4
21	59.5	12.7	217	6	US-10-454-437-82
22	59	12.6	111	7	US-11-072-512-2820
23	59	12.6	124	6	US-10-771-257-19
24	59	12.6	124	7	US-11-127-677-19
25	59	12.6	182	7	US-11-170-653-49

```

26 59 12.6 208 7 US-11-214-413-32 Sequence 32, Appl
27 59 12.6 381 7 US-11-087-099-3485 Sequence 3485, Ap
28 59 12.6 498 7 US-11-131-479-76 Sequence 76, Appl
29 58.5 12.5 211 7 US-11-170-653-50 Sequence 50, Appl
30 58.5 12.5 228 7 US-11-170-653-39 Sequence 39, Appl
31 58.5 12.5 248 7 US-11-054-515-1790 Sequence 1790, Ap
32 58.5 12.5 327 7 US-11-064-246-15 Sequence 15, Appl
33 58.5 12.5 327 7 US-11-064-246-16 Sequence 16, Appl
34 58.5 12.5 345 7 US-11-107-028-37 Sequence 37, Appl
35 58 12.4 121 7 US-11-107-028-35 Sequence 35, Appl
36 58 12.4 130 7 US-11-188-187A-3 Sequence 3, Appli
37 58 12.4 140 7 US-11-183-205-62 Sequence 62, Appl
38 58 12.4 378 7 US-11-087-099-8872 Sequence 8872, Ap
39 58 12.4 378 7 US-11-087-099-11183 Sequence 1183, A
40 58 12.4 391 7 US-11-207-626A-32 Sequence 32, Appl
41 58 12.4 451 7 US-11-124-620-5 Sequence 5, Appli
42 58 12.4 451 7 US-11-124-620-7 Sequence 7, Appli
43 58 12.4 464 7 US-11-087-099-7240 Sequence 7240, Ap
44 58 12.4 2204 7 US-11-052-554A-134 Sequence 134, App
45 57.5 12.3 108 7 US-11-097-812-39 Sequence 39, Appl

```

ALIGNMENTS

```

RESULT 1
US-10-517-939-328
; Sequence 328, Application US/10517939
; Publication No. US20060003433A1
; GENERAL INFORMATION:
; APPLICANT: Steer, Brian
; APPLICANT: Callen, Walter
; APPLICANT: Healey, Shaun
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Wu, Di
; APPLICANT: Blum, David
; APPLICANT: Esteghlalian, Alireza
; TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM
; TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 564462007901
; CURRENT APPLICATION NUMBER: US/10/517,939
; CURRENT FILING DATE: 2004-12-13
; PRIOR APPLICATION NUMBER: PCT/US03/19153
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/389,299
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 380
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 328
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample.
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)....(29)
US-10-517-939-328

```

```

Query Match 34.8%; Score 163; DB 6; Length 499;
Best Local Similarity 35.3%; Pred. No. 9.3e-12;
Matches 30; Conservative 21; Mismatches 30; Indels 4; Gaps. 2;

OY 5 VKQYKNDSPAGDQKQKGLQVNTGSSVLDSTVTVYWF-TRDGGSTLVNCDMAA 63
Db 354 LSVQVRTGDSVNSQIRQIHVKNSKTVALKNVTYVYNTKNGQN---FDCDYAK 410

OY 64 MCGNIRASFGSVNPATPTADTYLQ 88
Db 411 IGCSNVTHKFVTLQRPVKGADAYLE 435

RESULT 2

```

US-10-510-386-230
; Sequence 230, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510.386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 230
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-230

Query Match 33.5%; Score 157; DB 6; Length 534;
Best Local Similarity 37.6%; Pred. No. 5.2e-11;
Matches 32; Conservative 16; Mismatches 33; Indels 4; Gaps 2;

Qy 5 VKVQYKND SAPGDNQIKPGLQLVNTGSSVDLSTVTYRYWF-TRDGGSTLVYNCDAAM 63
Db IAVQYRAGDNNVNGNQIRPOLKNNKSKTIVSLNRITVRYWYKTRKQGN---FDCDYAQ 445

Qy 64 MGCNIRASFGSVNPATPTADTYIQ 88
Db IGCSTHFKFVLKAVNGADTYLE 470

RESULT 3
US-11-179-411-22
; Sequence 22, Application US/11179411
; Publication No. US20050266491A1
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: Szczepaniak, William
; TITLE OF INVENTION: RENILLA RENIFORMIS FLOURESCENT PROTEINS, NUCLEIC ACIDS ENCODING THEM
; TITLE OF INVENTION: FLOURESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIGH TH
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS
; FILE REFERENCE: 24729-128
; CURRENT APPLICATION NUMBER: US/11/179,411
; CURRENT FILING DATE: 2005-07-12
; PRIOR APPLICATION NUMBER: US/09/808,898
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/189,691
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 09/277,716
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 08/757,046
; PRIOR FILING DATE: 1996-11-25
; PRIOR APPLICATION NUMBER: 08/597,274
; PRIOR FILING DATE: 1996-02-06
; PRIOR APPLICATION NUMBER: 08/908,909
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 08/990,103
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Gaussia
US-11-179-411-22

Query Match 24.1%; Score 113; DB 7; Length 382;
Best Local Similarity 30.9%; Pred. No. 6.2e-06;
Matches 29; Conservative 21; Mismatches 32; Indels 12; Gaps 4;

Qy 5 VKVQYKND SAPGDNQIKPGLQLVNTGSSVDLSTVTYRYWFTRDGGSTLVYNCDAAM 64
Db MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKRYVYTS DGTQGTQTFW-CDHAGA 59

Qy 65 GCGN-----IRASF--GSVNPATPTADTYIQ 88
Db LLGNSYVDNTSKVTANFVKETASP-TSTYDTTYVE 92

RESULT 4
US-11-175-766-22
; Sequence 22, Application US/11175766
; Publication No. US20050272111A1
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: Szczepaniak, William
; TITLE OF INVENTION: RENILLA RENIFORMIS FLOURESCENT PROTEINS, NUCLEIC ACIDS ENCODING TH
; TITLE OF INVENTION: FLOURESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIGH TH
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS
; FILE REFERENCE: 24729-128
; CURRENT APPLICATION NUMBER: US/11/175,766
; CURRENT FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: US/09/808,898
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/189,691
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 09/277,716
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 08/757,046
; PRIOR FILING DATE: 1996-11-25
; PRIOR APPLICATION NUMBER: 08/597,274
; PRIOR FILING DATE: 1996-02-06
; PRIOR APPLICATION NUMBER: 08/908,909
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 08/990,103
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Gaussia
US-11-175-766-22

Query Match 24.1%; Score 113; DB 7; Length 382;
Best Local Similarity 30.9%; Pred. No. 6.2e-06;
Matches 29; Conservative 21; Mismatches 32; Indels 12; Gaps 4;

Qy 5 VKVQYKND SAPGDNQIKPGLQLVNTGSSVDLSTVTYRYWFTRDGGSTLVYNCDAAM 64
Db MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKRYVYTS DGTQGTQTFW-CDHAGA 59

Qy 65 GCGN-----IRASF--GSVNPATPTADTYIQ 88
Db LLGNSYVDNTSKVTANFVKETASP-TSTYDTTYVE 92

RESULT 5
US-10-517-939-344
; Sequence 344, Application US/10517939
; Publication No. US20060003433A1
; GENERAL INFORMATION:
; APPLICANT: Steer, Brian
; APPLICANT: Callen, Walter
; APPLICANT: Healey, Shaun
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Wu, Di
; APPLICANT: Blum, David
; APPLICANT: Esteghlalian, Alireza
; TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM
; TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM

; FILE REFERENCE: 564462007901
; CURRENT APPLICATION NUMBER: US/10/517,939
; CURRENT FILING DATE: 2004-12-13
; PRIOR APPLICATION NUMBER: PCT/US03/19153
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/389,299
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 380
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 344
; LENGTH: 738
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample.
US-10-517-939-344

Query Match 23.0%; Score 108; DB 6; Length 738;
Best Local Similarity 32.1%; Pred. No. 5.2e-05;
Matches 27; Conservative 18; Mismatches 31; Indels 8; Gaps 4;
QY 7 VQYNDSAPGDN--QIKPGLQLVNTGSSVDLSTVTVRYWFTRDGSGSTLVNCDWAAAM 64
DB 591 LQYR---SADGNNYQMKPQETIKNAGKVPILSELIIRYVTFPE-STQPDVTRIDWAO 646
QY 65 GCGNIRASFGSVNPTATPDATYLO 88
DB 647 GAHVQITV--VPPSDDAAAHAYVE 668

RESULT 6
US-11-170-653-51
; Sequence 51, Application US/11/170653
; Publication No. US20050271769A1
; GENERAL INFORMATION:
; APPLICANT: Danisco A/S
; APPLICANT: Sibbesen, Ole
; APPLICANT: Sorensen, Jens
; TITLE OF INVENTION: Xylanase Variants Having Altered Sensitivity to Xylanase Inhibitor
; FILE REFERENCE: 674509-2046
; CURRENT APPLICATION NUMBER: US/11/170,653
; CURRENT FILING DATE: 2005-06-23
; PRIOR APPLICATION NUMBER: US/10/237,386
; PRIOR FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: PCT/IB01/00426
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: GB 0005585.5
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: GB 0015751.1
; PRIOR FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 51
; LENGTH: 210
; TYPE: PRT
; ORGANISM: A. tubigenis
US-11-170-653-51

Query Match 15.4%; Score 72; DB 7; Length 210;
Best Local Similarity 31.5%; Pred. No. 0.23;
Matches 23; Conservative 10; Mismatches 22; Indels 18; Gaps 4;
QY 29 NTGSSVDLSTVTV-RVWFTDGDG-----GSSTLVYNCDAWACGNIASFG 74
DB 40 NLGFTYDESAGTFSMYW--EDGVSDPFVGLWGTGSSITTSYSAEYASGSASLYAVYG 97
QY 75 SVNPTATPDATYL 87
DB 98 WVN--YFQAEYII 108

RESULT 7
US-11-052-554A-281

; Sequence 281, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 281
; LENGTH: 5291
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-281

Query Match 14.9%; Score 70; DB 7; Length 5291;
Best Local Similarity 33.3%; Pred. No. 18;
Matches 29; Conservative 15; Mismatches 19; Indels 24; Gaps 7;
QY 17 GD---NQIKPGLQLVNTGSS--VDL-STVTVRYWFTRDG---SSTLVYNCDAW----- 61
DB 848 GDDVINAVEKGETLVVSGSTSGVEAGQTVTVTF-----GKKNYTTTVEANGSWTVNVPPA 902
QY 62 --AM--GCGNIRASFGSVNPTATPD 84
DB 903 DLALPDGAGNVQASVNSNGNSAQAD 929

RESULT 8
US-11-072-512-2825
; Sequence 2825, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOUYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2825
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-2825

```

Query Match      14.8%; Score 68.5; DB 7; Length 500;
Best Local Similarity 33.3%; Pred. No. 1.7;
Matches 19; Conservative 9; Mismatches 18; Indels 11; Gaps 3;

QY      31  GSSGVDLSTVTVRYWFRDQ-----GSS--TLVNCDD---WAAMGCGNIRASFGSV 76
      DB      414  GIPETDIDSCSRYTKADGTQCPSSGSGTTVIYTCEFISAYGARGSANIKVTFISV 470

RESULT 9
US-10-517-939-330
; Sequence 330, Application US/10517939
; Publication No. US20060003433A1
; GENERAL INFORMATION:
; APPLICANT: Steer, Brian
; APPLICANT: Callen, Walter
; APPLICANT: Healey, Shaun
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Wu, Di
; APPLICANT: Blum, David
; APPLICANT: Esteghlalian, Alireza
; TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM
; TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 564462007901
; CURRENT APPLICATION NUMBER: US/10/517,939
; CURRENT FILING DATE: 2004-12-13
; PRIOR APPLICATION NUMBER: PCT/US03/19153
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/389,299
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 380
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 330
; LENGTH: 755
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample.
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(35)
US-10-517-939-330

Query Match      14.5%; Score 68; DB 6; Length 755;
Best Local Similarity 30.1%; Pred. No. 3.1;
Matches 31; Conservative 8; Mismatches 40; Indels 24; Gaps 5;

QY      2  SGGVKQVYK-NNDSAPGDNQIKPGLQVNTGSSVDLSTTVRYWFRD----- 49
      DB      651  SGACTVAYAITNDWGSQ---FTANVTLTNTGSA--LNGWTLAYAFPNGQNTISNMGTA 705

QY      50  --GGSGTLVYNCDAAMGCGNIRASFG-----SVNEATPTADT 85
      DB      706  VOGSGSVSVTNAGWNGSLPPNVASFGFQASVSGNNSVPASFT 748

RESULT 10
US-11-207-626A-26
; Sequence 26, Application US/11207626A
; Publication No. US20060014276A1
; GENERAL INFORMATION:
; APPLICANT: Havenga, Menzo
; APPLICANT: Vogels, Ronald
; APPLICANT: Bout, Abraham
; TITLE OF INVENTION: CHIMERIC ADENOVIRUSES
; FILE REFERENCE: 2578-4123.2US
; CURRENT APPLICATION NUMBER: US/11/207,626A
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: EP 98202297.2
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26

```



```
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 380
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
US-10-517-939-14

Query Match      13.0%; Score 61; DB 6; Length 350;
Best Local Similarity 26.7%; Pred. No. 8.7;
Matches 16; Conservative 10; Mismatches 26; Indels 8; Gaps 1;

Qy 10 KNDSPAGDNOIKPGLQVNTGSSVDLSTVTYVRYFTRDGGSTLTVNCDW 61
Db 40 RHYNSLTADNQMKPESVLDRTATLAKGDLHLAAVDFTRVDALMYFARDNGIPMRVHTLAW 99

RESULT 13
US-11-102-476-46
; Sequence 46, Application US/11102476
; Publication No. US20050271680A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Palaniappan, Raghavan U.M.
; APPLICANT: Cornell Research Foundation, Inc.
; TITLE OF INVENTION: Novel Immunogenic Proteins of Leptospira
; FILE REFERENCE: 1153.080US1
; CURRENT APPLICATION NUMBER: US/11/102,476
; CURRENT FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: PCT/US2003/32385
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/417721
; PRIOR FILING DATE: 2002-10-10
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 1889
; TYPE: PRT
; ORGANISM: Leptospira interrogans
US-11-102-476-46

Query Match      13.0%; Score 61; DB 7; Length 1889;
Best Local Similarity 26.2%; Pred. No. 62;
Matches 21; Conservative 14; Mismatches 27; Indels 18; Gaps 4;

Qy 19 NOIKPGL-----QLVNTG-----SSVDLSTVTYVRYFTRDGGSTLTVY-----NCDWA 62
Db 498 NPVNFSLAKGLTKFTATGIYSDNSKNDITSVIT--WFSSDSSIATISNAKQNGNAYGA 555

Qy 63 AMGCNTRASFSGSVNPATPT 82
Db 556 ATGTTDIKATFGKVSVPST 575

RESULT 14
US-11-098-686-10278
; Sequence 10278, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10278
; LENGTH: 806
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10278

Query Match      12.9%; Score 60.5; DB 7; Length 806;
Best Local Similarity 33.3%; Pred. No. 26;
Matches 20; Conservative 5; Mismatches 20; Indels 15; Gaps 4;

Qy 10 KNDSPAGDNOIKPGLQVNTGSSVDLSTVTYVRYF--TRDG--GSSTLTVY--NCDWAA 63
Db 286 KNNIVTPGKDPVPSSQLVRNHQ-----RYWVRGMADGFIGNSTWCYISNLDYVS 336

RESULT 15
US-11-087-099-7483
; Sequence 7483, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 7483
; LENGTH: 983
; TYPE: PRT
; ORGANISM: Pseudomonas syringae pv. syringae B728a
US-11-087-099-7483

Query Match      12.9%; Score 60.5; DB 7; Length 983;
Best Local Similarity 33.3%; Pred. No. 33;
Matches 15; Conservative 9; Mismatches 10; Indels 11; Gaps 2;

Qy 2 SGGVKVQYKNDNSAPGD--NQIKPGLQL-----VNTGSSSV 35
Db 522 NAGVKAIEYFSNTSFSGDPLVTRVEFGVNLNWTGNSVTNAGSTAV 566

Search completed: March 2, 2006, 14:28:14
Job time : 2.8047 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 2, 2006, 12:06:58 ; Search time 32.1046 Seconds
(without alignments)
4927.737 Million cell updates/sec

Title: US-09-917-376-4

Perfect score: 469

Sequence: 1 VSGGVKQYKNDSPGDNQ.....RASFGSVNPATPTADTYLQX 89

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US09017376/runat_02032006_091454_8101/app_query.fasta_1
-DB=Issued_Patents_NA -QWTF=fastcap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abes06h -USER=US09017376 @CGN 1 278 @runat_02032006_091454_8101
-NCPU=6 -ICPU=3 -NO_WMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCUTS COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	195.5	41.7	2029	3	US-09-136-574A-46
2	194.5	41.5	6416	3	US-09-136-574A-2
3	192.5	41.0	11707	3	US-09-136-574A-1
4	182.5	38.9	2977	2	US-07-862-588B-1
5	165	35.2	1438	3	US-09-339-159B-3
6	165	35.2	1482	3	US-09-198-956-9
7	165	35.2	1482	3	US-09-198-955A-11
8	165	35.2	1482	3	US-09-694-531-11
9	165	35.2	1482	3	US-09-670-141-9

10	165	35.2	1482	3	US-10-072-152-11	Sequence 11, Appl
11	155.5	33.2	4059	3	US-09-784-554B-1	Sequence 1, Appli
12	150.5	32.1	4056	3	US-09-784-554B-3	Sequence 3, Appli
13	120.5	25.7	1624	2	US-07-862-588B-6	Sequence 5, Appli
14	120.5	25.7	1775	2	US-07-862-588B-5	Sequence 1, Appli
15	114	24.3	486	2	US-08-048-164A-1	Sequence 1, Appli
16	114	24.3	486	2	US-08-048-164A-3	Sequence 3, Appli
17	114	24.3	486	2	US-08-460-462-1	Sequence 1, Appli
18	114	24.3	486	2	US-08-460-462-3	Sequence 3, Appli
19	114	24.3	486	2	US-08-460-457-1	Sequence 1, Appli
20	114	24.3	486	2	US-08-460-457-3	Sequence 3, Appli
21	114	24.3	486	2	US-08-460-458-1	Sequence 1, Appli
22	114	24.3	486	2	US-08-460-458-3	Sequence 3, Appli
23	114	24.3	486	2	US-08-460-455-1	Sequence 1, Appli
24	114	24.3	486	2	US-08-460-455-3	Sequence 3, Appli
25	114	24.3	486	2	US-08-330-394A-1	Sequence 1, Appli
26	114	24.3	486	2	US-08-330-394A-3	Sequence 3, Appli
27	114	24.3	499	3	US-09-006-636-5	Sequence 5, Appli
28	114	24.3	499	3	US-09-006-632-5	Sequence 5, Appli
29	114	24.3	499	3	US-09-325-274-5	Sequence 5, Appli
30	113	24.1	1146	3	US-09-277-716-21	Sequence 21, Appl
31	113	24.1	1146	3	US-09-609-161B-21	Sequence 21, Appl
32	78.5	16.7	764	3	US-09-533-559-5404	Sequence 5404, Ap
33	77	16.4	8107	3	US-09-335-586-3	Sequence 3, Appli
34	76	16.2	136058	3	US-09-949-018-12565	Sequence 12565, A
35	76	16.2	136480	3	US-09-949-016-17064	Sequence 17064, A
36	75.5	16.1	5714	3	US-09-620-312D-393	Sequence 393, App
37	72.5	15.5	819	3	US-09-902-540-3752	Sequence 3752, Ap
38	72.5	15.5	17897	3	US-09-902-540-1182	Sequence 1182, Ap
39	72.5	15.5	20284	3	US-09-526-193A-21	Sequence 21, Appl
40	70	14.9	10892	3	US-09-902-540-962	Sequence 962, App
41	70	14.9	25165	3	US-09-453-702B-39	Sequence 39, Appl
42	70	14.9	25165	3	US-10-114-170-39	Sequence 39, Appl
43	70	14.9	246444	3	US-09-949-016-13113	Sequence 13113, A
44	69.5	14.8	103987	3	US-09-949-016-12513	Sequence 12513, A
45	69.5	14.8	103988	3	US-09-949-016-17050	Sequence 17050, A

ALIGNMENTS

RESULT 1

US-09-136-574A-46
; Sequence 46, Application US/09136574A
; Patent No. 6294366
; GENERAL INFORMATION:
; APPLICANT: Farrington, Graham K.
; Anderson, Paige
; Gibbs, Moreland
; Bergquist, Peter
; Daniels, Roy
; Morgan, Hugh W.
; Williams, Diane P.

TITLE OF INVENTION: Compositions and Methods for
Treating Cellulose Containing
Cellulase Enzyme Compositions

NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Center, P.O. Box 457

CITY: Spring House

STATE: PA

COUNTRY: USA

ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/136,574A

FILING DATE: 19-Aug-1998

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/932,571

Fabrics Using Truncated

```
; FILING DATE: September 19, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: 1997US001/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2029 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-136-574A-46

Alignment Scores:
Pred. No.: 5,09e-17 Length: 2029
Score: 195.50 Matches: 38
Percent Similarity: 60.2% Conservative: 15
Best Local Similarity: 43.2% Mismatches: 34
Query Match: 41.7% Indels: 1
DB: 3 Gaps: 1

US-09-917-376-4 (1-89) x US-09-136-574A-46 (1-2029)
Qy 1 ValSerGlyValGlnTyrLeuValAsnAspSerAlaProGlyAspAsnGln 20
Db 1 ATGGGAAGTGTGTGAAGGTACTCTACAGAACAAATGAGACAGTGGCAGCACAGGTTCT 60
Qy 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 61 ATAAGCGCGTGTAGATAGTGAATGGAGGACAGCAGCAGTGTGATCTTAGCAGGTT 120
Qy 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
Db 121 AAGTAAGACTACTGTTACACAGTGGATGGTCAACAGCCACAGAGTGGCGTA--TGTGAC 177
Qy 61 TrpAlaAlaMetGlyCysGlyValAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
Db 178 TGGGCACAGATAGGGGCAAGCAATGTGACATTCATTTTGTGAAGCTTAGCAGCGGAGTG 237
Qy 81 ProThrAlaAspThrTyrLeuGln 88
Db 238 AGTGGAGCGGATATTACCTGGAG 261

RESULT 2
US-09-136-574A-2
; Sequence 2, Application US/09136574A
; Patent No. 6294366
; GENERAL INFORMATION:
; APPLICANT: Farrington, Graham K.
; Anderson, Paige
; Gibbs, Moreland
; Bergquist, Peter
; Daniels, Roy
; Morgan, Hugh W.
; Williams, Diane P.
; TITLE OF INVENTION: Compositions and Methods for
; Treating Cellulose Containing Fabrics Using Truncated
; Cellulase Enzyme Compositions
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Center, P.O. Box 457
; CITY: Spring House
; STATE: PA
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
```

```
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/136,574A
; FILING DATE: 19-Aug-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/932,571
; FILING DATE: September 19, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: 1997US001/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6416 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-136-574A-2

Alignment Scores:
Pred. No.: 3,72e-16 Length: 6416
Score: 194.50 Matches: 38
Percent Similarity: 61.2% Conservative: 14
Best Local Similarity: 44.7% Mismatches: 32
Query Match: 41.5% Indels: 1
DB: 3 Gaps: 1

US-09-917-376-4 (1-89) x US-09-136-574A-2 (1-6416)
Qy 4 GlyValLysValGlnTyrLeuValAsnAspSerAlaProGlyAspAsnGlnIleLysPro 23
Db 2665 GGTGTGAAGGTACTCTACAGAACAAATGAGACAGTGGCAGCACAGGTTCTATAGGCCG 2724
Qy 24 GlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArg 43
Db 2725 TGGTTTAAAGATAGTGAATGGAGGACAGCAGCAGTGTGATCTTAGCAGGCTTAAGATAAGA 2784
Qy 44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
Db 2785 TACTGGTACACAGTGGATGGTGCAACAGCCACAGAGTGGCGTA--TGTGACTGGGCACAG 2841
Qy 64 MetGlyCysGlyValAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
Db 2842 ATAGGGCAACAAATGTGACATTCATTTTGTGAAGCTTAGCAGCGGAGTGGAGCG 2901
Qy 84 AspThrTyrLeuGln 88
Db 2902 GATTATTACCTGGAG 2916

RESULT 3
US-09-136-574A-1
; Sequence 1, Application US/09136574A
; Patent No. 6294366
; GENERAL INFORMATION:
; APPLICANT: Farrington, Graham K.
; Anderson, Paige
; Gibbs, Moreland
; Bergquist, Peter
; Daniels, Roy
; Morgan, Hugh W.
; Williams, Diane P.
; TITLE OF INVENTION: Compositions and Methods for
; Treating Cellulose Containing Fabrics Using Truncated
; Cellulase Enzyme Compositions
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
```

```

;
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Center, P.O. Box 457
; CITY: Spring House
; STATE: PA
; COUNTRY: USA
; ZIP: 19477
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/136,574A
; FILING DATE: 19-Aug-1998
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/932,571
; FILING DATE: September 19, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: 1997US001/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; TELEX: <Unknown>
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11707 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-136-574A-1

Alignment Scores:
Pred. No.: 1.69e-15 Length: 11707
Score: 192.50 Matches: 37
Percent Similarity: 62.4% Conservative: 16
Best Local Similarity: 43.5% Mismatches: 31
Query Match: 41.0% Indels: 1
DB: 3 Gaps: 1

US-09-917-376-4 (1-89) x US-09-136-574A-1 (1-11707)
QY 4 GlyValValGlnTyrIleAsnAspSerAlaProGlyAspAsnGlnIleLysPro 23
Db 4038 GGTITGAAGGTACTATACAAAGAACAAATGAGCAAGTTCCTAAGGCGG 4097

QY 24 GlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArg 43
Db 4098 TGGTTTAAGATAGTAAGTGGAGGCGAGCAGCAGTGTGATCTAGCAGGTTAAGATAAGA 4157

QY 44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
Db 4158 TACTGGTACACAGTGGATGTTGACAAAGCCACAGAGTGGCGTA--TGTGACTGGGCACAG 4214

QY 64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
Db 4215 ATAGGGGCAACAAATGTACATTCATTTTGTGAAGCTGAGCAGCGGAGTGAGTGGAGCG 4274

QY 84 AspThrTyrLeuGln 88
Db 4275 GATTATTACTTGGAG 4289

RESULT 4
US-07-862-588B-1
; Sequence 1, Application US/07862588B
; Patent No. 5916796
; GENERAL INFORMATION:
; APPLICANT: Joergensen, Per Linea
; APPLICANT: Sch lein, Martin
; APPLICANT: Hansen, Christian
; TITLE OF INVENTION: An Enzyme Exhibiting Cellulase Activity
;
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5916796o No. 5916796disk of No. 5916796th America, Inc.
; STREET: 405 Lexington Avenue, 62nd floor
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10017
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/862,588B
; FILING DATE: 19920727
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 164/90
; FILING DATE: 19-JAN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/DK91/00013
; FILING DATE: 18-JAN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Zelson, Steve T. / Lambiris, Elias J.
; REGISTRATION NUMBER: 30,335 / 33,728
; REFERENCE/DOCKET NUMBER: 3425.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2977 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Bacillus lautus
; STRAIN: NCIMB 40250
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 677..2776
; OTHER INFORMATION:
;
; US-07-862-588B-1
Alignment Scores:
Pred. No.: 5.76e-15 Length: 2977
Score: 182.50 Matches: 36
Percent Similarity: 66.3% Conservative: 19
Best Local Similarity: 43.4% Mismatches: 25
Query Match: 38.9% Indels: 3
DB: 2 Gaps: 2

US-09-917-376-4 (1-89) x US-07-862-588B-1 (1-2977)
QY 6 LysValGlnTyrIleAsnAsnAspSerAlaProGlyAspAsnGlnIleLysProGlyLeu 25
Db 144 CGACTGCAGTACAGCGCGCGCATACAAATGACGCGCACCAACCATGATCAAGCGCTCCTTC 203

QY 26 GlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArgTyrTrp 45
Db 204 AACATCAAAACAAACACGCTCTCGGCTGTGATTGATTAAGCAGCTCAAAATCCGCTACTAC 263

QY 46 PheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAlaMetGly 65
Db 264 TTCACCAAGATGGTTCCTCGGCGGTGAACGGTGG---ATCGACTGGCGCGCAGCTCGGC 320

QY 66 CysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAlaAspThr 85
Db 321 GGCAGCAACATTCAGATCTCGTTTGGC-----AACCATACTGCGCAGCAATTCGATACG 374

QY 86 TyrLeuGln 88
```


TYPE: DNA
ORGANISM: Clostridium thermocellum
US-09-198-955A-11

Alignment Scores:
Pred. No.: 5,78e-13 Length: 1482
Score: 165.00 Matches: 35
Percent Similarity: 59.8% Conservative: 23
Best Local Similarity: 36.1% Mismatches: 29
Query Match: 35.2% Indels: 10
DB: 3 Gaps: 2

US-09-917-376-4 (1-89) x US-09-198-955A-11 (1-1482)

QY 1 ValSerGlyGlyValGlnTyrLysAsnAspSerAlaProGlyAspAsnGln 20
DB 991 GTATCAGCAATTGAAGGTTGAATTTCTACACAGCAATCCTTCAGATACTACTACTCA 1050
QY 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
DB 1051 ATCAATCCTCAGTTCAAGGTTACTAATACCGAAGCAGTCGCAATTGATTTGTCCAACTC 1110
QY 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
DB 1111 ACATTGAGATATTATTATACAGTAGACGACAGAAAGATCAGACCTTCTGG---TGTGAC 1167
QY 61 TrpAlaAlaMet-----GlyCysGlyAsnIleArgAla 71
DB 1168 CATGCTGCAATAATCGGCAGTAACGGCAGCTACAACGGAATTACTTCAAAATGTAAGA 1227
QY 72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
DB 1228 ACATTGTGAAAAATGAGTTCTCTCAACAATAACGCAGACACCTACTCTTGAA 1278

RESULT 8

US-09-694-531-11
Sequence 11, Application US/09694531
Patent No. 6368843
GENERAL INFORMATION:
APPLICANT: Andersen, Lene N.
APPLICANT: Schulein, Martin
APPLICANT: Lange, Niels E.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Moller, Soren
APPLICANT: Glad, Sanne O. S.
APPLICANT: Kauppinen, Markus S.
APPLICANT: Schnorr, Kirk
APPLICANT: Kongsbak, Lars
TITLE OF INVENTION: No. 6368843el Pectate Lyases
FILE REFERENCE: 5378.200-US
CURRENT APPLICATION NUMBER: US/09/694,531
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 09/198,955
PRIOR FILING DATE: 1998-11-24
PRIOR APPLICATION NUMBER: 1343/97
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 1344/97
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/067,249
PRIOR FILING DATE: 1997-12-02
PRIOR APPLICATION NUMBER: 60/067,240
PRIOR FILING DATE: 1997-12-02
PRIOR APPLICATION NUMBER: 09/073,684
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 09/184,217
PRIOR FILING DATE: 1998-11-02
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 1482
TYPE: DNA
ORGANISM: Clostridium thermocellum
US-09-694-531-11

Alignment Scores:

Pred. No.: 5,78e-13 Length: 1482
Score: 165.00 Matches: 35
Percent Similarity: 59.8% Conservative: 23
Best Local Similarity: 36.1% Mismatches: 29
Query Match: 35.2% Indels: 10
DB: 3 Gaps: 2

US-09-917-376-4 (1-89) x US-09-694-531-11 (1-1482)

QY 1 ValSerGlyGlyValGlnTyrLysAsnAspSerAlaProGlyAspAsnGln 20
DB 991 GTATCAGCAATTGAAGGTTGAATTTCTACACAGCAATCCTTCAGATACTACTACTCA 1050
QY 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
DB 1051 ATCAATCCTCAGTTCAAGGTTACTAATACCGAAGCAGTCGCAATTGATTTGTCCAACTC 1110
QY 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
DB 1111 ACATTGAGATATTATTATACAGTAGACGACAGAAAGATCAGACCTTCTGG---TGTGAC 1167
QY 61 TrpAlaAlaMet-----GlyCysGlyAsnIleArgAla 71
DB 1168 CATGCTGCAATAATCGGCAGTAACGGCAGCTACAACGGAATTACTTCAAAATGTAAGA 1227
QY 72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
DB 1228 ACATTGTGAAAAATGAGTTCTCTCAACAATAACGCAGACACCTACTCTTGAA 1278

RESULT 9

US-09-670-141-9
Sequence 9, Application US/09670141
Patent No. 6429000
GENERAL INFORMATION:
APPLICANT: Andersen, Lene N.
APPLICANT: Schulein, Martin
APPLICANT: Lange, Niels Erik K.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Schnorr, Kirk
TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus
FILE REFERENCE: 5377.200-US
CURRENT APPLICATION NUMBER: US/09/670,141
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: 09/198,956
PRIOR FILING DATE: 1998-11-24
PRIOR APPLICATION NUMBER: 1344/97
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/067,240
PRIOR FILING DATE: 1997-12-02
NUMBER OF SEQ ID NOS: 26
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 9
LENGTH: 1482
TYPE: DNA
ORGANISM: Bacillus licheniformis
US-09-670-141-9

Alignment Scores:

Pred. No.: 5,78e-13 Length: 1482
Score: 165.00 Matches: 35
Percent Similarity: 59.8% Conservative: 23
Best Local Similarity: 36.1% Mismatches: 29
Query Match: 35.2% Indels: 10
DB: 3 Gaps: 2

US-09-917-376-4 (1-89) x US-09-670-141-9 (1-1482)

QY 1 ValSerGlyGlyValGlnTyrLysAsnAspSerAlaProGlyAspAsnGln 20
DB 991 GTATCAGCAATTGAAGGTTGAATTTCTACACAGCAATCCTTCAGATACTACTACTCA 1050

```

Qy 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 1051 ATCAATCCTCAGTTTACTAATACCGAAGCAGTGCATTTGTTTCCAAACTC 1110
Qy 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
Db 1111 ACATTGAGATATTATTATACATAGCAGCAGAAAGATCAGACCTTCTGG---TGTGAC 1167
Qy 61 TrpAlaAlaMet-----GlyCysGlyAsnIleArgAla 71
Db 1168 CATGCTGCAATAAATCGGCAGTAAACGGCAGCTACACGGAATTACTTCAAAATGTAAAGGA 1227
Qy 72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
Db 1228 ACATTGTGAAAAATGAGTTCTCTCAACAAATAACGACAGACCTTCTGAA 1278

RESULT 10
US-09-917-376-4 (1-89) x US-10-072-152-11 (1-1482)
; Sequence 11, Application US/10072152
; Patent No. 6677147
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schuilein, Martin
; APPLICANT: Lange, Niels E.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Moller, Soren
; APPLICANT: Glad, Sanne O. S.
; APPLICANT: Kauppinen, Markku S.
; APPLICANT: Schnorr, Kirk
; APPLICANT: Kongsbak, Lars
; TITLE OF INVENTION: No. 6677147el Pectate Lyases
; FILE REFERENCE: 5378.200-US
; CURRENT APPLICATION NUMBER: US/10/072,152
; CURRENT FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: US/09/198,955
; PRIOR FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: 1343/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 1344/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/067,249
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 60/067,240
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 09/073,684
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 09/184,217
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1482
; TYPE: DNA
; ORGANISM: Clostridium thermocellum
US-10-072-152-11

Alignment Scores:
Pred. No.: 5.78e-13 Length: 1482
Score: 165.00 Matches: 35
Percent Similarity: 59.8% Conservative: 23
Best Local Similarity: 36.1% Mismatches: 29
Query Match: 35.2% Indels: 10
DB: 3 Gaps: 2

US-09-917-376-4 (1-89) x US-10-072-152-11 (1-1482)
Qy 1 ValSerGlyGlyValGlnTyrLysAsnAspSerAlaProGlyAspAsnGln 20
Db 991 GTATCAGCAATTGAGGTTGATTTCTACACAGCAATCTTCAGATACCTAACTCA 1050
Qy 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 1051 ATCAATCCTCAGTTTACTAATACCGAAGCAGTGCATTTGTTTCCAAACTC 1110

```

```

Db 1051 ATCAATCCTCAGTTTACTAATACCGAAGCAGTGCATTTGTTTCCAAACTC 1110
Qy 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
Db 1111 ACATTGAGATATTATTATACATAGCAGCAGAAAGATCAGACCTTCTGG---TGTGAC 1167
Qy 61 TrpAlaAlaMet-----GlyCysGlyAsnIleArgAla 71
Db 1168 CATGCTGCAATAAATCGGCAGTAAACGGCAGCTACACGGAATTACTTCAAAATGTAAAGGA 1227
Qy 72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
Db 1228 ACATTGTGAAAAATGAGTTCTCTCAACAAATAACGACAGACCTTCTGAA 1278

RESULT 11
US-09-784-554B-1
; Sequence 1, Application US/09784554B
; Patent No. 6815192
; GENERAL INFORMATION:
; APPLICANT: Schnorr, Kirk
; APPLICANT: Jorgensen, Per Lina
; APPLICANT: Schuilein, Martin
; TITLE OF INVENTION: FAMILY 44 XYLOGLYCANASES
; FILE REFERENCE: 10017.200-US
; CURRENT APPLICATION NUMBER: US/09/784,554B
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4059
; TYPE: DNA
; ORGANISM: Paenibacillus polymyxa
US-09-784-554B-1

Alignment Scores:
Pred. No.: 5.25e-11 Length: 4059
Score: 155.50 Matches: 27
Percent Similarity: 56.5% Conservative: 21
Best Local Similarity: 31.8% Mismatches: 36
Query Match: 33.2% Indels: 1
DB: 3 Gaps: 1

US-09-917-376-4 (1-89) x US-09-784-554B-1 (1-4059)
Qy 4 GlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnIleLysPro 23
Db 3598 GGACTGTGCTCCCAATATCGCACAGCATACTTAAGGTGAACAGCAATACCTCAATCCG 3657
Qy 24 GlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArg 43
Db 3658 CAATTCCAAATTGTAAACAAGGACACCTCCATACCGATCAACGAGTTGAAATTCGC 3717
Qy 44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
Db 3718 TACTACTACAAATCGACGGTGACCGTGAGCAG---ACTTTCACACTGCGACTATGCGACG 3774
Qy 64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
Db 3775 CTGAGCTGCTCAAGCTGAACGGTAACTGGTTAAATGAGAGAGCTGCAACCGGTGCC 3834
Qy 84 AspThrTyrLeuGln 88
Db 3835 GATTATTATTGGA 3849

RESULT 12
US-09-784-554B-3
; Sequence 3, Application US/09784554B
; Patent No. 6815192
; GENERAL INFORMATION:
; APPLICANT: Schnorr, Kirk
; APPLICANT: Jorgensen, Per Lina
; APPLICANT: Schuilein, Martin
; TITLE OF INVENTION: FAMILY 44 XYLOGLYCANASES

```

```
; FILE REFERENCE: 10017-200-US
; CURRENT APPLICATION NUMBER: US/09/784,554B
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 4056
; TYPE: DNA
; ORGANISM: Paenibacillus polymyxa
US-09-784-554B-3

Alignment Scores:
Pred. No.: 2,61e-10 Length: 4056
Score: 150.50 Matches: 27
Percent Similarity: 56.5% Conservative: 21
Best Local Similarity: 31.8% Mismatches: 36
Query Match: 32.1% Indels: 1
DB: 3 Gaps: 1

US-09-917-376-4 (1-89) x US-09-784-554B-3 (1-4056)
QY 4 GlyVallyValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnIleLysPro 23
DB 3595 CGATTGGTCTCCAGTATCGCACGGGATCAAAATGTGAACGACATCATTGACCCG 3654
QY 24 GlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArg 43
DB 3655 CATTTCCAAATTTAAATAAAGGTACATCTCCGTACCGATCAACGAGTGAATAATTCG 3714
QY 44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaIa 63
DB 3715 TACTACTACAGCATCGACGGTGACCGTGCAGCAG---ACATTCAACTGCGACTATCGCGTG 3771
QY 64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
DB 3772 CTGAGCTGCTCGAAGCTGAATGTAAGTGTAAATGGATAAGCTGCAACCGGTGCT 3831
QY 84 AspThrTyrLeuGln 88
DB 3832 GATTATTATTGGAA 3846

RESULT 13
US-07-862-588B-6
; Sequence 6, Application US/07862588B
; Patent No. 5916796
; GENERAL INFORMATION:
; APPLICANT: Joergensen, Per Linea
; APPLICANT: Sch lein, Martin
; APPLICANT: Hansen, Christian
; TITLE OF INVENTION: An Enzyme Exhibiting Cellulase Activity
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 59167960 No. 5916796disk of No. 5916796th America, Inc.
; STREET: 405 Lexington Avenue, 62nd floor
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/862,588B
; FILING DATE: 19920727
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 164/90
; FILING DATE: 19-JAN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/DK91/00013
; FILING DATE: 18-JAN-1991

; ATTORNEY/AGENT INFORMATION:
; NAME: Zelson, Steve T. / Lambiris, Elias J.
; REGISTRATION NUMBER: 30,335 / 33,728
; REFERENCE/DOCKET NUMBER: 3425.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1624 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Bacillus lautus
; STRAIN: NCIMB 40250
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 30..1607
; OTHER INFORMATION:
US-07-862-588B-6

Alignment Scores:
Pred. No.: 1,06e-06 Length: 1624
Score: 120.50 Matches: 28
Percent Similarity: 54.1% Conservative: 18
Best Local Similarity: 32.9% Mismatches: 36
Query Match: 25.7% Indels: 3
DB: 2 Gaps: 2

US-09-917-376-4 (1-89) x US-07-862-588B-6 (1-1624)
QY 2 SerGlyGlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnIle 21
DB 1323 ACGGGAACTTGTGTTCATACAAAGTTGGCGACACTAGCCCGGATACCAATG 1382
QY 22 LysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThr 41
DB 1383 AAGCTTCTTTAAACATCAAGAACACCGGTACAAACCCCTGTTAACTGACGGCGCTCAAG 1442
QY 42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 61
DB 1443 CTNNNNNNNNNNNNNAAAAGAC---GGACCTGGCGGATATGAGCTGCTCGATCGACTGG 1499
QY 62 AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro 81
DB 1500 GCGCAATCGGCCGGAACGAATGTTCTGCTGGCATTG-----GCTAACTTTTACCGGAGT 1553
QY 82 ThrAlaAspThrTyr 86
DB 1554 AATACGGATACCTTAC 1568

RESULT 14
US-07-862-588B-5
; Sequence 5, Application US/07862588B
; Patent No. 5916796
; GENERAL INFORMATION:
; APPLICANT: Joergensen, Per Linea
; APPLICANT: Sch lein, Martin
; APPLICANT: Hansen, Christian
; TITLE OF INVENTION: An Enzyme Exhibiting Cellulase Activity
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 59167960 No. 5916796disk of No. 5916796th America, Inc.
; STREET: 405 Lexington Avenue, 62nd floor
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```



```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/862,588B
; FILING DATE: 19920727
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 164/90
; FILING DATE: 19-JAN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/DK91/00013
; FILING DATE: 18-JAN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Zelson, Steve T. / Lambiris, Elias J.
; REGISTRATION NUMBER: 30,335 / 33,728
; REFERENCE/DOCKET NUMBER: 3425.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1775 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Bacillus lautus
; STRAIN: NCIMB 40250
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 30..(1625..1775)
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /partial
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /transl_except= (pos: 1446 .. 1458, aa:
; OTHER INFORMATION: OTR)
US-07-862-588B-5

```

```

Alignment Scores:
Pred. No.: 1.2e-06 Length: 1775
Score: 120.50 Matches: 28
Percent Similarity: 54.1% Conservative: 18
Best Local Similarity: 32.9% Mismatches: 36
Query Match: 25.7% Indels: 3
DB: 2 Gaps: 2

```

```

US-09-917-376-4 (1-89) x US-07-862-588B-5 (1-1775)
QY 2 SerGlyGlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnIle 21
Db 1323 ACGGGGAACCTGTGTTCACAAATGTTGGGACACTACGCGCCACGGATAACCAATG 1382
QY 22 LysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThr 41
Db 1383 AAGCCTTCCTTTACATCAAGACACGCGTACACCCCTGTTACCTGAGCGGCTCAAG 1442
QY 42 ValArgTyrTrpPheThrArgAspGlyGlySerThrLeuValTyrAsnCysAspTrp 61
Db 1443 CTNNNNNNNNNNNNNNAAAAAGAC---GGACCTGGGATATGAGTCTCGATCGACTGG 1499
QY 62 AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro 81
Db 1500 GCCCAATCGCGCAAGCAATGTTCTCTGGCATTC-----GCTAACTTTACCGGGAGT 1553
QY 82 ThrAlaAspThrTyr 86
Db 1554 AATACGATACTTAC 1568
RESULT 15
US-08-048-164A-1

```

```

; Sequence 1, Application US/08048164A
; Patent No. 5496934
; GENERAL INFORMATION:
; APPLICANT: Shoseyov, Oded
; APPLICANT: Shpiegl, Itai
; APPLICANT: Goldstein, Marc A.
; APPLICANT: Doi, Roy H.
; TITLE OF INVENTION: CELLULOSE BINDING DOMAIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/048,164A
; FILING DATE: 14-APR-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7809-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 486 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..486
US-08-048-164A-1

```

```

Alignment Scores:
Pred. No.: 1.49e-06 Length: 486
Score: 114.00 Matches: 29
Percent Similarity: 52.6% Conservative: 22
Best Local Similarity: 29.9% Mismatches: 34
Query Match: 24.3% Indels: 12
DB: 2 Gaps: 4

```

```

US-09-917-376-4 (1-89) x US-08-048-164A-1 (1-486)
QY 2 SerGlyGlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnIle 21
Db 7 ACATCATCAATGTCAGTTGAATTTTACAACTCTTACAAATCAGCAACAACTCAAT 66
QY 22 LysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThr 41
Db 67 ACACCAATATCAAAATTTACTAACACATCTCACAGTGTATTAATTTAAATGACGTAAA 126
QY 42 ValArgTyrTrpPheThrArgAspGlyGlySerThrLeuValTyrAsnCysAspTrp 61
Db 127 GTTAGATATTATTACACAAGTGATGTGTACACAGGACAACTTTCTGG---TGTGACCAT 183
QY 62 AlaAlaMetGlyCysGlyAsn-----IleArgAlaSerPhe 73
Db 184 GCTGTGTCATTTATTAGGAAATAGCTGTGTGATAACTAGCAAGTAGCAACAACTTC 243
QY 74 -----GlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88

```

Db 244 GTTAAAGAAACAGCAAGCCCA--ACATCAACCTATGATACATATGTTGAA 291

Search completed: March 2, 2006, 13:50:30
Job time : 38.1046 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 2, 2006, 12:19:26 ; Search time 167.029 Seconds
(without alignments)
4406.259 Million cell updates/sec

Title: US-09-917-376-4

Perfect score: 469

Sequence: 1 VSGGVKQYKNDSPAGDNQ.....RASFGSVNPATPTADTYLQX 89

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O=/abs/ABSSWEB.spool/US0917376/runat_02032006_091500_8223/app_query.fasta_1
-DB=Published_Applications_NA_Main -QFMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=BLOSUM62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs06p
-USER=US0917376 @CGN 1 1 3095 @runat_02032006_091500_8223 -NCPUP=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA_Main:

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3:	/cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4:	/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5:	/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6:	/cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7:	/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8:	/cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9:	/cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10:	/cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	469	100.0	2289	3	US-09-917-378-2
2	466	99.4	3687	3	US-09-917-384-2
3	466	99.4	3687	3	US-09-917-383-2
4	462	98.5	2869	3	US-09-917-376-2
5	462	98.5	2869	5	US-10-155-400-2
6	257	54.8	2223	6	US-10-156-761-550
7	257	54.8	9025608	6	US-10-156-761-1

8	190	40.5	1527	6	US-10-369-493-46838
9	188	40.1	1434	8	US-10-466-208-11
10	188	40.1	1491	8	US-10-466-208-7
11	188	40.1	2510	7	US-10-433-577-34
12	165	35.2	501	9	US-10-933-404-1
13	165	35.2	1438	6	US-10-372-054-3
14	165	35.2	1482	5	US-10-072-152-11
15	165	35.2	1482	7	US-10-655-433-11
16	157	33.5	1314	3	US-09-974-300-684
17	155.5	33.2	4059	3	US-09-784-554B-1
18	155.5	33.2	4059	3	US-10-896-555-1
19	150.5	32.1	4056	3	US-09-784-554B-3
20	150.5	32.1	4056	8	US-10-896-555-3
21	131.5	28.0	3180	9	US-10-886-393A-29
22	130.5	27.8	4452	7	US-10-282-122A-15299
23	126.5	27.0	1303	9	US-10-490-737-1
24	126.5	27.0	1747	9	US-10-490-737-2
25	126.5	27.0	3489	9	US-10-490-737-3
26	114	24.3	768	6	US-10-261-446-5
27	114	24.3	768	7	US-10-261-445B-5
28	114	24.3	768	8	US-10-782-234-5
29	113	24.1	652	8	US-10-748-055-9
30	113	24.1	1146	3	US-09-808-898-21
31	78.5	16.7	764	8	US-10-653-047-5404
32	78.5	16.7	3147	7	US-10-282-122A-13716
33	76	16.2	757	7	US-10-767-701-8547
34	76	16.2	96597	7	US-10-052-482-226
35	75.5	16.1	1041	7	US-10-260-238-742
36	75.5	16.1	1041	7	US-10-437-963-76751
37	75.5	16.1	2091	7	US-10-211-462-222
38	75.5	16.1	2091	8	US-10-723-860-4038
39	75.5	16.1	2091	8	US-10-756-149-4006
40	75.5	16.1	3396	7	US-10-398-458-1
41	75.5	16.1	4041	3	US-09-747-835A-21
42	75.5	16.1	4041	7	US-10-312-312-21
43	75.5	16.1	4353	5	US-10-225-567A-405
44	75.5	16.1	4755	9	US-10-505-486-201
45	75.5	16.1	4914	3	US-09-974-298-188

ALIGNMENTS

RESULT 1

US-09-917-378-2
; Sequence 2, Application US/0917378
; Publication No. US20030119093A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT MANNANASE FROM ACIDOTHERMUS
; FILE REFERENCE: 40197.7US01
; CURRENT APPLICATION NUMBER: US/09/917,378
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2289
; TYPE: DNA
; ORGANISM: Acidothermus cellulolyticus
US-09-917-378-2

Alignment Scores:
Pred. No.: 6.94e-56 Length: 2289
Score: 469.00 Matches: 88
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-09-917-376-4 (1-89) x US-09-917-378-2 (1-2289)

Qy 1 ValSerGlyVallyValGlnTyrLysAsnAspSerAlaProGlyAspAsnGln 20
Db 1363 GTGTCGGGTGGGTGAAGTGCAGTACAGAACAAATGATTCGGCGCGGTGATAACAG 1422
Qy 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 1423 ATCAACCGGGTCTCAGTTGGTGAATACCGGGTGGTGGTGGATTTGTTCGACGGTG 1482
Qy 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
Db 1483 ACGTGGGTACTGGTTCACCGGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1542
Qy 61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
Db 1543 TGGCGCGCATGGGTGTGGGAATATCCGCGCTCGTTCGGTGGTGAACCGCGGACG 1602
Qy 81 ProThrAlaAspThrTyrLeuGln 88
Db 1603 CCGACGGCGACACCTACCTGCAG 1626

RESULT 2

US-09-917-384-2

; Sequence 2, Application US/09917384

; Publication No. US20030096342A1

; GENERAL INFORMATION:

; APPLICANT: DING, SHI-YOU

; APPLICANT: ADNEY, WILLIAM S.

; APPLICANT: VINZANT, TODD B.

; APPLICANT: DECKER, STEPHEN R.

; APPLICANT: HIMMEL, MICHAEL E.

; TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE FROM ACIDOTHERMUS

; FILE REFERENCE: 40170.6US01

; CURRENT APPLICATION NUMBER: US/09/917,384

; CURRENT FILING DATE: 2001-07-28

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 2

; LENGTH: 3687

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Segment of

; OTHER INFORMATION: Guxa

US-09-917-384-2

Alignment Scores:

Pred. No.: 3.5e-55 Length: 3687
Score: 466.00 Matches: 87
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 98.9% Mismatches: 0
Query Match: 99.4% Indels: 0
DB: 3 Gaps: 0

US-09-917-376-4 (1-89) x US-09-917-384-2 (1-3687)

Qy 1 ValSerGlyVallyValGlnTyrLysAsnAspSerAlaProGlyAspAsnGln 20
Db 1750 GTGTCGGGTGGGTGAAGTGCAGTACAGAACAAATGATTCGGCGCGGTGATAACAG 1809
Qy 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 1810 ATCAACCGGGTCTCAGTTGGTGAATACCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1869
Qy 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
Db 1870 ACGTGGGTACTGGTTCACCGGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1929
Qy 61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
Db 1930 TGGCGCGCATGGGTGTGGGAATATCCGCGCTCGTTCGGTGGTGAACCGCGGACG 1989

Qy 81 ProThrAlaAspThrTyrLeuGln 88
Db 1990 CCGACGGCGGACACCTACCTGCAG 2013

RESULT 3

US-09-917-383-2

; Sequence 2, Application US/09917383

; Publication No. US20030104522A1

; GENERAL INFORMATION:

; APPLICANT: DING, SHI-YOU

; APPLICANT: ADNEY, WILLIAM S.

; APPLICANT: VINZANT, TODD B.

; APPLICANT: DECKER, STEPHEN R.

; APPLICANT: HIMMEL, MICHAEL E.

; TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE FROM ACIDOTHERMUS

; FILE REFERENCE: 40170.6US01

; CURRENT APPLICATION NUMBER: US/09/917,383

; CURRENT FILING DATE: 2001-07-28

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 2

; LENGTH: 3687

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Segment of

; OTHER INFORMATION: Guxa

US-09-917-383-2

Alignment Scores:

Pred. No.: 3.5e-55 Length: 3687
Score: 466.00 Matches: 87
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 98.9% Mismatches: 0
Query Match: 99.4% Indels: 0
DB: 3 Gaps: 0

US-09-917-376-4 (1-89) x US-09-917-383-2 (1-3687)

Qy 1 ValSerGlyVallyValGlnTyrLysAsnAspSerAlaProGlyAspAsnGln 20
Db 1750 GTGTCGGGTGGGTGAAGTGCAGTACAGAACAAATGATTCGGCGCGGTGATAACAG 1809
Qy 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 1810 ATCAACCGGGTCTCAGTTGGTGAATACCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1869
Qy 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
Db 1870 ACGTGGGTACTGGTTCACCGGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1929
Qy 61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
Db 1930 TGGCGCGCATGGGTGTGGGAATATCCGCGCTCGTTCGGTGGTGAACCGCGGACG 1989
Qy 81 ProThrAlaAspThrTyrLeuGln 88
Db 1990 CCGACGGCGGACACCTACCTGCAG 2013

RESULT 4

US-09-917-376-2

; Sequence 2, Application US/09917376

; Publication No. US20040038334A1

; GENERAL INFORMATION:

; APPLICANT: DING, SHI-YOU

; APPLICANT: ADNEY, WILLIAM S.

; APPLICANT: VINZANT, TODD B.

; APPLICANT: HIMMEL, MICHAEL E.

; TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS

; FILE REFERENCE: 40197.4US01

```

; CURRENT APPLICATION NUMBER: US/09/917,376
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2869
; TYPE: DNA
; ORGANISM: Acidothermus cellulolyticus
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (2869)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-917-376-2
Alignment Scores:
Pred. No.: 9,34e-55 Length: 2869
Score: 462.00 Matches: 86
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 97.7% Mismatches: 0
Query Match: 98.5% Indels: 0
DB: 3 Gaps: 0

US-09-917-376-4 (1-89) x US-09-917-376-2 (1-2869)
QY 1 ValSerGlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGln 20
Db 2605 GTGTCGGGTGGGGTGAAGGTGCAGTATAAGATAATGATTCGGCGCGGGTGATAATCAG 2664
QY 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 2665 ATCAAGCCGGGTTCGAGGTGGTGAATACCGGGTCGTGCGTGGATTTGTCACCGGTG 2724
QY 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
Db 2725 ACGGTGGGTACTGTTTACCCGGGATGGTGGCTCGTGCACACTGGTGACACTGTGAC 2784
QY 61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
Db 2785 TGGCGCGCGATCGGGTGTGGGAATATCCGGCCTCGTTCGGTCCGTGAACCCGCGGACG 2844
QY 81 ProThrAlaAspThrTyrLeuGln 88
Db 2845 CCGACGGCGGACACTACCTGCAG 2868

RESULT 5
US-10-155-400-2
; Sequence 2, Application US/10155400
; Publication No. US2003010898A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
; FILE OF INVENTION: CELLULOYTICUS
; FILE REFERENCE: NREL 01-36A
; CURRENT APPLICATION NUMBER: US/10/155,400
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2869
; TYPE: DNA
; ORGANISM: Acidothermus cellulolyticus
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (2869)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-10-155-400-2
Alignment Scores:
Pred. No.: 9,34e-55 Length: 2869
Score: 462.00 Matches: 86

US-09-917-376-4 (1-89) x US-10-155-400-2 (1-2869)
QY 1 ValSerGlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGln 20
Db 2605 GTGTCGGGTGGGGTGAAGGTGCAGTATAAGATAATGATTCGGCGCGGGTGATAATCAG 2664
QY 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 2665 ATCAAGCCGGGTTCGAGGTGGTGAATACCGGGTCGTGCGTGGATTTGTCACCGGTG 2724
QY 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
Db 2725 ACGGTGGGTACTGTTTACCCGGGATGGTGGCTCGTGCACACTGGTGACACTGTGAC 2784
QY 61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
Db 2785 TGGCGCGCGATCGGGTGTGGGAATATCCGGCCTCGTTCGGTCCGTGAACCCGCGGACG 2844
QY 81 ProThrAlaAspThrTyrLeuGln 88
Db 2845 CCGACGGCGGACACTACCTGCAG 2868

RESULT 6
US-10-156-761-550
; Sequence 550, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 550
; LENGTH: 2223
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2223)
US-10-156-761-550
Alignment Scores:
Pred. No.: 1.17e-25 Length: 2223
Score: 257.00 Matches: 49
Percent Similarity: 71.3% Conservative: 13
Best Local Similarity: 56.3% Mismatches: 25
Query Match: 54.8% Indels: 0
DB: 6 Gaps: 0

US-09-917-376-4 (1-89) x US-10-156-761-550 (1-2223)
QY 2 SerGlyGlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGln 21
Db 1771 TCGGGCGGCTCAAGGTCTCTACAGAACAACTCTCTCGGCCACCGAACCGCATC 1830
QY 22 LysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThr 41
Db 22 LysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThr 41
```

Db 706456 GGAGCCGACGCCTACCTCGAA 706476

RESULT 8

US-10-369-493-46838

; Sequence 46838, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 46838

; LENGTH: 1527

; TYPE: DNA

; ORGANISM: Bacillus subtilis

; US-10-369-493-46838

Alignment Scores:

Pred. No.:	2,55e-16	Length:	1527
Score:	190.00	Matches:	35
Percent Similarity:	65.1%	Conservative:	21
Best Local Similarity:	40.7%	Mismatches:	26
Query Match:	40.5%	Indels:	4
DB:	6	Gaps:	2

US-09-917-376-4 (1-89) x US-10-369-493-46838 (1-1527)

QY	4	GlyVallysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGlnIleLysPro	23
Db	1084	GGTATTCTGTACAGTACAGAGCGGGATGGAGCTATGACAGCAACCAATCCGTCGG	1143
QY	24	GlyLeuGlnIleValAsnThrGlySerSerSerValAspLeuSerThrValThrValArg	43
Db	1144	CAGCTTCAATAAAAAATAACGGCAATACACCGTTGATTTAAAGATGCTACTGCCGT	1203
QY	44	TyrTrpPhe--ThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTyrAla	62
Db	1204	TACTGGTATTAAGCGCAAAACAAAGGCCAAAC-----TTTGACTGTGACTACGGC	1254
QY	63	AlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThr	82
Db	1255	CAGATTGGATCGCGCAATGTGCACACACAAGTTTGTGACGTTGCATAAACCAAGCAAGT	1314
QY	83	AlaAspThrTyrLeuGln 88	
Db	1315	GCAGATACCTATCTGGAA 1332	

RESULT 9

US-10-466-208-11

; Sequence 11, Application US/10466208

; Publication No. US20040180348A1

; GENERAL INFORMATION:

; APPLICANT: GENOFOCUS Co., Ltd.

; TITLE OF INVENTION: Method for Surface Display of Proteins on Genetic Carriers

; FILE REFERENCE: 3260-13

; CURRENT APPLICATION NUMBER: US/10/466,208

; CURRENT FILING DATE: 2003-07-14

; PRIOR APPLICATION NUMBER: KR2001-2156

; PRIOR FILING DATE: 2001-01-15

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: Kopatentcin 1.71

; SEQ ID NO 11

; LENGTH: 1434

; TYPE: DNA

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CMCase gene with additional his encoding sequences
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1431)
US-10-466-208-11
Alignment Scores:
Pred. No.: 4.53e-16 Length: 1434
Score: 188.00 Matches: 34
Percent Similarity: 61.2% Conservative: 18
Best Local Similarity: 40.0% Mismatches: 31
Query Match: 40.1% Indels: 2
DB: 8 Gaps: 1

US-09-917-376-4 (1-89) x US-10-466-208-11 (1-1434)
Qy 4 GlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnIleLysPro 23
Db 991 GGTATTCTGTACAATACAGACGAGGAGTGGAGTATGAACAGCAACCAATCCGTCG 1050
Qy 24 GlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArg 43
Db 1051 CAGCTTCAATAAATAAATAACGCAATACCAACCGTTGATTAAAGATGTCACGCGCGT 1110
Qy 44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
Db 1111 TACTGTATTAACGCGAATAAAGCCAAACGTT-----GACTGTGACTACGCGCAG 1164
Qy 64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
Db 1165 CTGGATCGGCAATGTGACATACAAGTTTGTGACGTTGCATATAAACCAAGCAAGGTGCA 1224
Qy 84 AspThrTyrLeuGln 88
Db 1225 GATACCTATCTGGAA 1239

RESULT 10
US-10-466-208-7
; Sequence 7, Application US/10466208
; Publication No. US20040180349A1
; GENERAL INFORMATION:
; APPLICANT: GENOFOCUS Co., Ltd.
; TITLE OF INVENTION: Method for Surface Display of Proteins on Genetic Carriers
; FILE REFERENCE: 3260-13
; CURRENT APPLICATION NUMBER: US/10/466,208
; CURRENT FILING DATE: 2003-07-14
; PRIOR APPLICATION NUMBER: KR2001-2156
; PRIOR FILING DATE: 2001-01-15
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Kopatentin 1.71
; SEQ ID NO 7
; LENGTH: 1491
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CMCase gene with mutated signal sequence to enhance its
; OTHER INFORMATION: hydrophobicity
; NAME/KEY: CDS
; LOCATION: (1)..(1488)
US-10-466-208-7
Alignment Scores:
Pred. No.: 4.77e-16 Length: 1491
Score: 188.00 Matches: 34
Percent Similarity: 61.2% Conservative: 18
Best Local Similarity: 40.0% Mismatches: 31
Query Match: 40.1% Indels: 2
DB: 8 Gaps: 1

US-09-917-376-4 (1-89) x US-10-466-208-7 (1-1491)
Qy 4 GlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnIleLysPro 23
Db 991 GGTATTCTGTACAATACAGACGAGGAGTGGAGTATGAACAGCAACCAATCCGTCG 1050
Qy 24 GlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArg 43
Db 1051 CAGCTTCAATAAATAAATAACGCAATACCAACCGTTGATTAAAGATGTCACGCGCGT 1110
Qy 44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
Db 1111 TACTGTATTAACGCGAATAAAGCCAAACGTT-----GACTGTGACTACGCGCAG 1164
Qy 64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
Db 1165 CTGGATCGGCAATGTGACATACAAGTTTGTGACGTTGCATATAAACCAAGCAAGGTGCA 1224
Qy 84 AspThrTyrLeuGln 88
Db 1225 GATACCTATCTGGAA 1239

RESULT 11
US-10-433-577-34
; Sequence 34, Application US/10433577
; Publication No. US20040171065A1
; GENERAL INFORMATION:
; APPLICANT: Genofocus Co., Ltd.
; TITLE OF INVENTION: Method for Expression of Proteins on Spore Surface
; FILE REFERENCE: PCT-Genofocus-1
; CURRENT APPLICATION NUMBER: US/10/433,577
; CURRENT FILING DATE: 2003-06-05
; PRIOR APPLICATION NUMBER: KR2000-74835
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Kopatentin 1.71
; SEQ ID NO 34
; LENGTH: 2510
; TYPE: DNA
; ORGANISM: Bacillus subtilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (461)..(2491)
US-10-433-577-34
Alignment Scores:
Pred. No.: 9.54e-16 Length: 2510
Score: 188.00 Matches: 34
Percent Similarity: 61.2% Conservative: 18
Best Local Similarity: 40.0% Mismatches: 31
Query Match: 40.1% Indels: 2
DB: 7 Gaps: 1

US-09-917-376-4 (1-89) x US-10-433-577-34 (1-2510)
Qy 4 GlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnIleLysPro 23
Db 2051 GGTATTCTGTACAATACAGACGAGGAGTGGAGTATGAACAGCAACCAATCCGTCG 2110
Qy 24 GlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArg 43
Db 2111 CAGCTTCAATAAATAAATAACGCAATACCAACCGTTGATTAAAGATGTCACGCGCGT 2170
Qy 44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
Db 2171 TACTGTATTAACGCGAATAAAGCCAAACGTT-----GACTGTGACTACGCGCAG 2224
Qy 64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
Db 2225 CTGGATCGGCAATGTGACATACAAGTTTGTGACGTTGCATATAAACCAAGCAAGGTGCA 2284
Qy 84 AspThrTyrLeuGln 88
Db 2285 GATACCTATCTGGAA 2299

```

```
RESULT 12
US-10-933-404-1
; Sequence 1, Application US/10933404
; Publication No. US20050118729A1
; GENERAL INFORMATION:
; APPLICANT: MORAG, Ely
; TITLE OF INVENTION: MICROARRAYS OF CELLULOSE BINDING CHIMERIC PROTEINS AND METHODS OF
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 85189-7100
; CURRENT FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: US/10/933,404
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: PCT/IL03/00177
; PRIOR FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US 60/362,061
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 501
; TYPE: DNA
; ORGANISM: Clostridium thermocellum
US-10-933-404-1
Alignment Scores:
Pred. No.: 2,13e-13 Length: 501
Score: 165.00 Matches: 35
Percent Similarity: 59.8% Conservatives: 23
Best Local Similarity: 36.1% Mismatches: 29
Query Match: 35.2% Indels: 10
DB: 2

US-09-917-376-4 (1-89) x US-10-933-404-1 (1-501)
Qy 1 ValSerGlyGlyValGlnTyrLysAsnAspSerAlaProGlyAspAsnGln 20
Db 13 GTATCAGCAATTGAAGTTGAATCTTACACAGCAATCTTCAGATACCTAACTCA 72
Qy 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 73 ATCAATCCTCAGTTCAGGTTACTTAACCGGAGCAGTCAATGTTGTTCCAACTC 132
Qy 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
Db 133 ACATTGAGATATTATTATACAGTAGACGACGACAGAAAGATCAGACCTTCTGG--TGTGAC 189
Qy 61 TrpAlaAlaMet-----GlyCysGlyAsnIleArgAla 71
Db 190 CATGCTGCAATAATCGGCAGTACGCGAGCTACAACGGAATTACTTCAAAATGTAAAGGA 249
Qy 72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
Db 250 ACATTGTAATAATGAGTTCTTCAACAAATAACGACGACACCTACCTTGAA 300

RESULT 13
US-10-372-054-3
; Sequence 3, Application US/10372054
; Publication No. US20030203466A1
; GENERAL INFORMATION:
; APPLICANT: Kauppinen, Markus
; APPLICANT: Schulein, Martin
; APPLICANT: Schnorr, Kirk
; APPLICANT: Andersen, Lene
; APPLICANT: Bjornvad, Mads
; TITLE OF INVENTION: Novel Mannanases
; FILE REFERENCE: 5440.204-US
; CURRENT APPLICATION NUMBER: US/10/372,054
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: US/09/339,159B
; PRIOR FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3

US-09-917-376-4 (1-89) x US-10-372-054-3 (1-1438)
Qy 1 ValSerGlyGlyValGlnTyrLysAsnAspSerAlaProGlyAspAsnGln 20
Db 940 GTATCAGCAATTGAAGTTGAATCTTACACAGCAATCTTCAGATACCTAACTCA 999
Qy 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 1000 ATCAATCCTCAGTTCAGGTTACTTAATACCGGAGCAGTCAATGTTGTTCCAACTC 1059
Qy 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
Db 1060 ACATTGAGATATTATTATACAGTAGACGACGACAGAAAGATCAGACCTTCTGG--TGTGAC 1116
Qy 61 TrpAlaAlaMet-----GlyCysGlyAsnIleArgAla 71
Db 1117 CATGCTGCAATAATCGGCAGTACGCGAGCTACAACGGAATTACTTCAAAATGTAAAGGA 1176
Qy 72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
Db 1177 ACATTGTAATAATGAGTTCTTCAACAAATAACGACGACACCTACCTTGAA 1227

RESULT 14
US-10-072-152-11
; Sequence 11, Application US/10072152
; Publication No. US20020142438A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schulein, Martin
; APPLICANT: Lange, Niels E.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Moller, Soren
; APPLICANT: Glad, Samne O. S.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Schnorr, Kirk
; APPLICANT: Kongsbak, Lars
; TITLE OF INVENTION: No. US20020142438A1el Pectate Lyases
; FILE REFERENCE: 5378.200-US
; CURRENT APPLICATION NUMBER: US/10/072,152
; CURRENT FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: US/09/198,955
; PRIOR FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: 1343/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 1344/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/067,249
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 60/067,240
; PRIOR APPLICATION NUMBER: 1997-12-02
; PRIOR APPLICATION NUMBER: 09/073,684
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 09/184,217
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1482
; TYPE: DNA
; ORGANISM: Clostridium thermocellum
```


US-10-072-152-11

Alignment Scores:
Pred. No.: 9,04e-13 Length: 1482
Score: 165.00 Matches: 35
Percent Similarity: 59.8% Conservative: 23
Best Local Similarity: 36.1% Mismatches: 29
Query Match: 35.2% Indels: 10
DB: 5 Gaps: 2

US-09-917-376-4 (1-89) x US-10-072-152-11 (1-1482)

QY 1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
DB 991 GTATCAGGCAATTTGAAGTTGAATCTACACAGCAATCCTTCAGATACTACTAATCA 1050
QY 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
DB 1051 ATCAATCTCAGTTCAAGTTACTAATACCGAAGCAGTGCAATTTGTTCCAAACTC 1110
QY 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
DB 1111 ACATTGAGATATTATTATACAGTACGACGACAGAAAGATCAGACCTTCTGG---TGTGAC 1167
QY 61 TrpAlaAlaMet-----GlyCysGlyAsnIleArgAla 71
DB 1168 CATGTCGAATAATCGGCAGTAACGCGAGTACACAGCAATTAATCTCAAAATGTAAAGGA 1227
QY 72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
DB 1228 ACATTGTGAAAAATGAGTTCTTCAACAAATAACGCGACACACCTTACCTTGAA 1278

RESULT 15

US-10-655-433-11
; Sequence 11, Application US/10655433
; Publication No. US20040067572A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schuelein, Martin
; APPLICANT: Lange, Niels E.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Moller, Soren
; APPLICANT: Glad, Sanne O. S.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Schnort, Kirk
; APPLICANT: Kongsbak, Lars
; TITLE OF INVENTION: No. US20040067572A1el Pectate Lyases
; FILE REFERENCE: 5378.200-US
; CURRENT APPLICATION NUMBER: US/10/655,433
; CURRENT FILING DATE: 2003-09-04
; PRIOR APPLICATION NUMBER: US/09/198,955A
; PRIOR FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: 1343/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 1344/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/067,249
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 60/067,240
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 09/073,684
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 09/184,217
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1482
; TYPE: DNA
; ORGANISM: Clostridium thermocellum
US-10-655-433-11

Alignment Scores:

Pred. No.: 9,04e-13 Length: 1482
Score: 165.00 Matches: 35
Percent Similarity: 59.8% Conservative: 23
Best Local Similarity: 36.1% Mismatches: 29
Query Match: 35.2% Indels: 10
DB: 7 Gaps: 2
US-09-917-376-4 (1-89) x US-10-655-433-11 (1-1482)
QY 1 ValSerGlyGlyValLysValGlnTyrLysAsnAsnAspSerAlaProGlyAspAsnGln 20
DB 991 GTATCAGGCAATTTGAAGTTGAATCTACACAGCAATCCTTCAGATACTACTAATCA 1050
QY 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
DB 1051 ATCAATCTCAGTTCAAGTTACTAATACCGAAGCAGTGCAATTTGTTCCAAACTC 1110
QY 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
DB 1111 ACATTGAGATATTATTATACAGTACGACGACAGAAAGATCAGACCTTCTGG---TGTGAC 1167
QY 61 TrpAlaAlaMet-----GlyCysGlyAsnIleArgAla 71
DB 1168 CATGTCGAATAATCGGCAGTAACGCGAGTACACAGCAATTAATCTCAAAATGTAAAGGA 1227
QY 72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
DB 1228 ACATTGTGAAAAATGAGTTCTTCAACAAATAACGCGACACACCTTACCTTGAA 1278
Search completed: March 2, 2006, 20:17:40
Job time : 697.029 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 2, 2006, 12:24:59 ; Search time 95.5064 Seconds
(without alignments)
2043.020 Million cell updates/sec

Title: US-09-917-376-4

Perfect score: 459

Sequence: 1 VSGGVKQVKNDSAPGDNO.....RASFGSVNRPATPDYVQLX 89

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 7218512 seqs, 1096188070 residues

Total number of hits satisfying chosen parameters: 14437024

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abs/ABSSWEB_spool/US09917376/runat_02032006_091502_8293/app_query.fasta.1
-DB=Published Applications NA_New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abes03p
-USER=US09917376 -CGN_1_1_1067@runat_02032006_091502_8293 -NCPU=6 -ICPU=3
-NO MWAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA_New:

1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2.*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3.*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq4.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	163	34.8	1500	8	US-10-517-939-327 Sequence 327, App
2	157	33.5	2602	8	US-10-510-386-229 Sequence 229, App
3	113	24.1	1146	12	US-11-179-411-21 Sequence 21, Appl
4	113	24.1	1146	12	US-11-175-766-21 Sequence 21, Appl

5	108	23.0	2217	8	US-10-517-939-343	Sequence 343, App
6	70	14.9	15876	12	US-11-052-554A-660	Sequence 660, App
7	69.5	14.8	617	6	US-09-925-065A-353393	Sequence 353393, A
8	68.5	14.6	1084	8	US-10-750-623-31182	Sequence 31182, A
9	68.5	14.6	1084	8	US-10-750-623-31182	Sequence 31182, A
10	68.5	14.6	2231	9	US-11-072-512-855	Sequence 855, App
11	68.5	14.6	4952	12	US-11-136-527-487	Sequence 487, App
12	68	14.5	2268	8	US-10-517-939-329	Sequence 329, App
13	68	14.5	35344	8	US-10-995-561-13307	Sequence 13307, A
14	67.5	14.4	556	6	US-09-925-065A-228751	Sequence 228751, A
15	67.5	14.4	2564	9	US-11-072-512-279	Sequence 279, App
16	67	14.3	523	6	US-09-925-065A-39541	Sequence 39541, A
17	67	14.3	171486	12	US-11-121-086-105	Sequence 105, App
18	66.5	14.2	600	8	US-10-750-185-1174	Sequence 1174, App
19	66.5	14.2	600	8	US-10-750-623-1174	Sequence 1174, App
20	66.5	14.2	1847	12	US-11-000-688-1099	Sequence 1099, App
21	66.5	14.2	2565	8	US-10-750-185-48986	Sequence 48986, A
22	66.5	14.2	2565	8	US-10-750-623-48986	Sequence 48986, A
23	65.5	14.0	617	6	US-09-925-065A-353394	Sequence 353394, A
24	65.5	14.0	1218	6	US-09-925-065A-547306	Sequence 547306, A
25	65.5	14.0	1218	6	US-09-925-065A-547307	Sequence 547307, A
26	65.5	14.0	1218	6	US-09-925-065A-547308	Sequence 547308, A
27	64.5	13.8	600	6	US-09-925-065A-336480	Sequence 336480, A
28	64.5	13.8	600	6	US-09-925-065A-336481	Sequence 336481, A
29	64.5	13.8	601	6	US-09-925-065A-729386	Sequence 729386, A
30	64.5	13.8	2370	6	US-09-925-065A-91998	Sequence 91998, A
31	64.5	13.8	2370	6	US-09-925-065A-91999	Sequence 91999, A
32	64.5	13.8	159138	8	US-10-995-561-13230	Sequence 13230, A
33	64	13.6	556	6	US-09-925-065A-369647	Sequence 369647, A
34	64	13.6	585	6	US-09-925-065A-385747	Sequence 385747, A
35	64	13.6	611	6	US-09-925-065A-658978	Sequence 658978, A
36	64	13.6	659	6	US-09-925-065A-87693	Sequence 87693, A
37	64	13.6	659	6	US-09-925-065A-87694	Sequence 87694, A
38	64	13.6	1898	9	US-11-072-512-1678	Sequence 1678, App
39	64	13.6	1990	12	US-11-062-225-2	Sequence 2, Appl
40	64	13.6	1990	12	US-11-062-225-10	Sequence 10, Appl
41	64	13.6	3074	12	US-11-000-688-315	Sequence 315, App
42	63.5	13.5	596	6	US-09-925-065A-521812	Sequence 521812, A
43	63.5	13.5	597	6	US-09-925-065A-258810	Sequence 258810, A
44	63.5	13.5	597	6	US-09-925-065A-258811	Sequence 258811, A
45	63.5	13.5	606	6	US-09-925-065A-484851	Sequence 484851, A

ALIGNMENTS

RESULT 1
US-10-517-939-327
; Sequence 327, Application US/10517939
; Publication No. US20060003433A1
; GENERAL INFORMATION:
; APPLICANT: Steer, Brian
; APPLICANT: Callen, Walter
; APPLICANT: Healey, Shaun
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Wu, Di
; APPLICANT: Blum, David
; APPLICANT: Esteghalian, Alireza
; TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM
; FILE REFERENCE: 564462007901
; CURRENT APPLICATION NUMBER: US/10/517,939
; CURRENT FILING DATE: 2004-12-13
; PRIOR APPLICATION NUMBER: PCT/US03/19153
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/389,299
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 380
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 327
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:

OTHER INFORMATION: Obtained from an environmental sample.
US-10-517-939-327

Alignment Scores:
Pred. No.: 8.87e-11 Length: 1500
Score: 163.00 Matches: 30
Percent Similarity: 60.0% Conservative: 21
Best Local Similarity: 35.3% Mismatches: 30
Query Match: 34.8% Indels: 4
DB: 8 Gaps: 2

US-09-917-376-4 (1-89) x US-10-517-939-327 (1-1500)

Qy 5 VallysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnLeuLysProGly 24
Db 1060 CTATCTGTGAATACAGACAGGGGAGTGTGAACACCAACCAATCCGCTCAG 1119
Qy 25 LeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArgTyr 44
Db 1120 ATCCATGTGAAACAAACACAGCAAGCACCCTTAATTTAAATAATGTAACCTCCGCTAC 1179
Qy 45 TrpPhe---ThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
Db 1180 TGGTATACACGAAACAAAGGCCAAAC-----TTGACTGTGACTACCGAAG 1230
Qy 64 MetGlyCysGlyAsnLeuArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
Db 1231 ATCGGATGACCAATGTGACCAAGTTTGTGACATTACAAACCTGTAACAAAGGTGCA 1290
Qy 84 AspThrTyrLeuGln 88
Db 1291 GATGCTATCTGGAA 1305

RESULT 2

US-10-510-386-229
Sequence 229, Application US/10510386
Publication No. US20050244922A1
GENERAL INFORMATION:
APPLICANT: Andersen, Jens Tonne
APPLICANT: Clausen, Ib Groth
APPLICANT: Jorgensen, Steen Troels
APPLICANT: Olsen, Peter Bjarke
APPLICANT: Rasmussen, Michael Dolberg
TITLE OF INVENTION: Improved Bacillus Host Cell
FILE REFERENCE: 10294.204-US
CURRENT APPLICATION NUMBER: US/10/510,386
CURRENT FILING DATE: 2004-10-04
NUMBER OF SEQ ID NOS: 248
SOFTWARE: PatentIn version 3.3
SEQ ID NO 229
LENGTH: 2602
TYPE: DNA
ORGANISM: Bacillus licheniformis
FEATURE:
NAME/KEY: CDS
LOCATION: (501)..(2102)

US-10-510-386-229

Alignment Scores:
Pred. No.: 9.84e-10 Length: 2602
Score: 157.00 Matches: 32
Percent Similarity: 56.5% Conservative: 16
Best Local Similarity: 37.6% Mismatches: 33
Query Match: 33.5% Indels: 4
DB: 8 Gaps: 2

US-09-917-376-4 (1-89) x US-10-510-386-229 (1-2602)

Qy 5 VallysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnLeuLysProGly 24
Db 1665 ATGACAGTACAAACAGCGGGGAGCAACAATGTAACGCGCAACCAATCCGCTCAG 1724
Qy 25 LeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArgTyr 44

Db 1725 CTCACATTAAACAAACAGCAAAAAACCGGTCTTTAAATCGAATCACTGTCCGCTAC 1784
Qy 45 TrpPhe---ThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
Db 1785 TGGTATAAACGAATCGCAAGGACAAAAT-----TTGACTGCGACTATGCCCAA 1835
Qy 64 MetGlyCysGlyAsnLeuArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
Db 1836 ATCGCTGCAGCAAAATCAGCACAATTCGTTCAATTAAAAAAGCGGTAAACGCGAGCA 1895
Qy 84 AspThrTyrLeuGln 88
Db 1896 GACACGTATCTGGAA 1910
RESULT 3
US-11-179-411-21
Sequence 21, Application US/11179411
Publication No. US20050266491A1
GENERAL INFORMATION:
APPLICANT: Bryan, Bruce
APPLICANT: Szent-Gyorgyi, Christopher
APPLICANT: Szczepaniak, William
TITLE OF INVENTION: RENILLA RENIFORMIS FLOURESCENT PROTEINS, NUCLEIC ACIDS ENCODING TI
TITLE OF INVENTION: FLOURESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIGH TH
FILE REFERENCE: 24729-128
CURRENT APPLICATION NUMBER: US/11/179,411
CURRENT FILING DATE: 2005-07-12
PRIOR APPLICATION NUMBER: US/09/808,898
PRIOR FILING DATE: 2001-03-15
PRIOR APPLICATION NUMBER: 60/189,691
PRIOR FILING DATE: 2000-03-15
PRIOR APPLICATION NUMBER: 09/277,716
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 08/757,046
PRIOR FILING DATE: 1996-11-25
PRIOR APPLICATION NUMBER: 08/597,274
PRIOR FILING DATE: 1996-02-06
PRIOR APPLICATION NUMBER: 08/908,909
PRIOR FILING DATE: 1997-08-08
PRIOR APPLICATION NUMBER: 08/990,103
PRIOR FILING DATE: 1997-12-12
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21
LENGTH: 1146
TYPE: DNA
ORGANISM: Gaussia
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1146)
OTHER INFORMATION: Nucleotide sequence encoding a CBD-Gaussia
OTHER INFORMATION: luciferase fusion protein
US-11-179-411-21
Alignment Scores:
Pred. No.: 0.00017 Length: 1146
Score: 113.00 Matches: 29
Percent Similarity: 53.2% Conservative: 21
Best Local Similarity: 30.9% Mismatches: 32
Query Match: 24.1% Indels: 12
DB: 12 Gaps: 4
US-09-917-376-4 (1-89) x US-11-179-411-21 (1-1146)
Qy 5 VallysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnLeuLysProGly 24
Db 1 ATGTCAGTTGAATTTTACAACTTAACAAATAACAGCAACAACTCAATTACACCAATA 60
Qy 25 LeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArgTyr 44
Db 61 ATCAAAATTACTTACACATCTGACAGTATTAAATTTAAATGAGTAAAGTTAGATAT 120

Db	1882	TATTTACGCCGGAG---	AGCAGCGCCGGTGGATAC	CAGGATCGACTGGGCCCCAATTC	1938
Qy	65	GlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAlaasp	84		
Db	1939	GGAGCAGAGCATGTC	CAGACGCGT	-----GTTCCGCGATCCGATGCGCGCGCGCAC	1992
Qy	85	ThrTyrLeuGln	88		
Db	1993	GCCTATGTCGAG	2004		
RESULT 6					

```
US-11-052-554A-660
; Sequence 660, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 660
; LENGTH: 15876
; TYPE: DNA
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-660

Alignment Scores:
Pred. No.: 1,15e+03 Length: 15876
Score: 70.00 Matches: 29
Percent Similarity: 50.6% Conservatives: 15
Best Local Similarity: 33.3% Mismatches: 19
Query Match: 14.9% Indels: 24
DB: 12 Gaps: 7

US-09-917-376-4 (1-89) x US-11-052-554A-660 (1-15876)
Qy 17 GlyAsp-----AsnGlnIleLysProGlyLeuGlnLeuValAsnThrGlySer 33
Db 2542 GGTGATGATGTGATCAACGGCGTCGAAAGGCGAAGCGTGTGCTGTAAGCGGACGACC 2601

Qy 34 Ser---ValAspLeu---SerThrValThrValArgTyrTrpPheThrArgAspGlyGly 51
Db 2602 AGCGGTGTGCAAGCGGCGGACGCGTACCGTTCACCTTT-----GGCGGC 2646

Qy 52 -----SerSerThrLeuValTyrAsnCysAspTrp----- 61
Db 2647 AAAAAATACACACACAGTGGAGTAAACCGGTAGCTGGACGGTGAATGTGCGCGCTGCC 2706

Qy 62 -----AlaAlaMet-----GlyCysGlyAsnIleArgAlaSerPheGlySerValAsn 77
Db 2707 GATCTCGTGGCTACCGGACGGCGGGCGGACGTGCGACGGAGTGTCACTAATATTAA 2766

Qy 78 ProAlaThrProThrAlaAsp 84
Db 2767 GGCAACAGCGGCCCGGCGCGAT 2787

RESULT 7
US-09-925-065A-353393
; Sequence 353393, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
```

```
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 353393
; LENGTH: 617
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-353393

Alignment Scores:
Pred. No.: 31.8 Length: 617
Score: 69.50 Matches: 19
Percent Similarity: 44.8% Conservatives: 11
Best Local Similarity: 28.4% Mismatches: 20
Query Match: 14.8% Indels: 17
DB: 6 Gaps: 2

US-09-917-376-4 (1-89) x US-09-925-065A-353393 (1-617)
Qy 25 LeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArgTyr 44
Db 214 TTGAGAGCAGTTAAGGGAGGGTCAAGATCTTGTAGCCTCAGCTCCATGACTCTAGGCCA 273

Qy 45 TrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla--- 63
Db 274 TGGTTT-----CTAATCTTATGCGCTAGTTTCTTGGTCTGCTCCC 312

Qy 64 -----MetGlyCysGlyAsnIleArgAlaSerPheGly 74
Db 313 CAGCAAGAGGTAAAGCCTATCTTAAAGAGGGCGCTGTACCATCTTTGTTTGTAGACTATAA 372

Qy 75 SerValAsnProAlaThrPro 81
Db 373 ACTATAAACCCAGCTCTCTCCA 393

RESULT 8
US-10-750-185-31182
; Sequence 31182, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31182
; LENGTH: 1084
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-31182

Alignment Scores:
Pred. No.: 81.8 Length: 1084
Score: 68.50 Matches: 25
Percent Similarity: 40.2% Conservatives: 12
Best Local Similarity: 27.2% Mismatches: 32
Query Match: 14.6% Indels: 23
DB: 8 Gaps: 5

US-09-917-376-4 (1-89) x US-10-750-185-31182 (1-1084)
Qy 16 ProGlyAspAsnGlnIleLysProGlyLeuGlnLeuVal-----AsnThr 30
Db 738 CCAGGCACCAAC-----CCCAATTCTCAGCTGGTCTCTCTCTCTCTCTCTCTCTG 788
```


; PRIOR FILING DATE: 2005-05-26
 ; NUMBER OF SEQ ID NOS: 362830
 ; SOFTWARE: Patent in version 3.2
 ; SEQ ID NO 487
 ; LENGTH: 4952
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 US-11-136-527-487

Alignment Scores:
 Pred. No.: 470 Length: 4952
 Score: 68.50 Matches: 22
 Percent Similarity: 44.6% Conservative: 7
 Best Local Similarity: 33.8% Mismatches: 25
 Query Match: 14.6% Indels: 11
 DB: 12 Gaps: 3

US-09-917-376-4 (1-89) x US-11-136-527-487 (1-4952)

Qy 31 GlySerSerValAspLeuSerThrValThrValArgTyrTrpPheThrArgAspGly 50
 Db 1413 GGGACCCAGAACAGACCTAGAGTCTAGCTGCGACGACTTATACACTCAAGGCGAGATGGA 1472
 Qy 51 -----GlySerSer-----ThrLeuValTyrAsnCysAsp----- 60
 Db 1473 ACCAGTGTCCAGTGGTCTCTCGAACACACAGCTATCTACACGTGTGAGTTGCTGCAGC 1532
 Qy 61 ---TipAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAla 79
 Db 1533 GTCTACGGACCAAGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1592
 Qy 80 ThrProThrAlaAsp 84
 Db 1593 ACAATAACCCCGGAC 1607

RESULT 12

US-10-517-939-329
 ; Sequence 329, Application US/10517939
 ; Publication No. US2006003433A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Steer, Brian
 ; APPLICANT: Callen, Walter
 ; APPLICANT: Healey, Shaun
 ; APPLICANT: Hazlewood, Geoff
 ; APPLICANT: Wu, Di
 ; APPLICANT: Blum, David
 ; APPLICANT: Esteghlalian, Alireza

; TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM
 ; FILE REFERENCE: 564462007901
 ; CURRENT APPLICATION NUMBER: US/10/517,939
 ; PRIORITY FILING DATE: 2004-12-13
 ; PRIOR APPLICATION NUMBER: PCT/US03/19153
 ; PRIOR FILING DATE: 2003-06-16
 ; PRIOR APPLICATION NUMBER: 60/389,299
 ; PRIOR FILING DATE: 2002-06-14
 ; NUMBER OF SEQ ID NOS: 380
 ; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 329
 ; LENGTH: 2268
 ; TYPE: DNA
 ; ORGANISM: Unknown

; FEATURE:
 ; OTHER INFORMATION: Obtained from an environmental sample.

US-10-517-939-329

Alignment Scores:
 Pred. No.: 222 Length: 2268
 Score: 68.00 Matches: 31
 Percent Similarity: 37.9% Conservative: 8
 Best Local Similarity: 30.1% Mismatches: 40
 Query Match: 14.5% Indels: 24
 DB: 8 Gaps: 5

US-09-917-376-4 (1-89) x US-10-517-939-329 (1-2268)

Qy 2 SerGlyGlyValIysValGlnTyrIys---AsnAsnAspSerAlaProGlyAspAsnGln 20
 Db 1951 TCAGGGCGCTGCACCGTGCCTACGCCATCACCACGACTGGGGCAGCGGT----- 2001
 Qy 21 IleIysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
 Db 2002 TTCACCGCCAAACGTATACCTCACCATACTGGCGGAAGCGCC-----CTCAACGGCTGG 2055
 Qy 41 ThrValArgTyrTrpPheThrArgAsp----- 49
 Db 2056 ACCCTGGCTATGCTTTCCCGCAATCAAAACCATCAGCAACGCTGGAAACGGAACGGCC 2115
 Qy 50 -----GlyGlySerSerThrLeuValTyrAsnCysAspTyrAlaAlaMetGlyCysGly 67
 Db 2116 GTTCAGTCCGCGACGAGCTCAGCGTCCACCAACGCGGTGGAAATGGCAGCTGCGCGCCC 2175
 Qy 68 AsnIleArgAlaSerPheGly-----SerValAsnProAlaThrProThr 82
 Db 2176 AACGTCTCCGCGACGCTTTGGCTTCCAGCGAGCTACAGCGGCAATAACAGCGTCCCTGCC 2235
 Qy 83 AlaAspThr 85
 Db 2236 AGCTTTACG 2244

RESULT 13

US-10-995-561-13307/c
 ; Sequence 13307, Application US/10995561
 ; Publication No. US20050272054A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
 ; FILE REFERENCE: CL001559
 ; CURRENT APPLICATION NUMBER: US/10/995,561
 ; CURRENT FILING DATE: 2004-11-24
 ; NUMBER OF SEQ ID NOS: 85702
 ; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13307
 ; LENGTH: 35344
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens

US-10-995-561-13307

Alignment Scores:
 Pred. No.: 5.23e+03 Length: 35344
 Score: 68.00 Matches: 16
 Percent Similarity: 57.1% Conservative: 4
 Best Local Similarity: 45.7% Mismatches: 11
 Query Match: 14.5% Indels: 4
 DB: 8 Gaps: 1

US-09-917-376-4 (1-89) x US-10-995-561-13307 (1-35344)

Qy 51 GlySerSerThrLeuValTyr-----AsnCysAspTyrAlaAlaMetGlyCys 66
 Db 23968 GGTAAGACCAACACTTCCCTATTCACCTTGGAGTCACTGAGCGCGCTTGTGGC 23909
 Qy 67 GlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro 81
 Db 23908 TTTAGAGCTAGGCGATCATTTACCACCTTCTTCCCTCTAGTCT 23864

RESULT 14

US-09-925-065A-228751/c
 ; Sequence 228751, Application US/09925065A
 ; Publication No. US20040181048A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single
 ; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome

```
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 228751
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-228751

Alignment Scores:
Pred. No.: 51      Length: 556
Score: 67.50      Matches: 25
Percent Similarity: 37.6%      Conservative: 7
Best Local Similarity: 29.4%      Mismatches: 20
Query Match: 14.4%      Indels: 33
DB: 6      Gaps: 4

US-09-917-376-4 (1-89) x US-09-925-065A-228751 (1-556)
QY 23 ProGlyLeuGlnLeuValAsnThrGlySer----- 32
Db 238 CCGTGGACTCCACCTTGTCATACAGGAGTTTAAGGAATGCTTTGCTGACAAAGACTTTT 179
QY 33 -----SerSerValAspLeuSerThrValThrValArgTyrTTPheThrArg 48
Db 178 ATTTAGCTTGTAGTTCTCTCGAATTTGTTTGTAGATTGAGTTACTGTGTAATAATATGG 119
QY 49 AspGlyGlySerSerThrLeuValTyrAsnCysAspTTPAlaAlaMet----- 64
Db 118 GACCGGGG-----TAGGACTGGTGGGGCTCTGCATCCCTTCC 80
QY 65 -----GlyCysGlyAsnIleArg---AlaSerPheGlySerVal 76
Db 79 TCCGAGGGGCGCTGGGAGAGGGTGTGGAGTGGCCCGCTGGTGTCTTCTCCCCCAATT 20
QY 77 AsnProAlaThrPro 81
Db 19 TCCTGTTGGACCCCT 5

RESULT 15
US-11-072-512-279/c
; Sequence 279, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTYUKI
; APPLICANT: NAGAHARI, KENJI

; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 279
; LENGTH: 2564
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-072-512-279

Alignment Scores:
Pred. No.: 296      Length: 2564
Score: 67.50      Matches: 23
Percent Similarity: 42.1%      Conservative: 9
Best Local Similarity: 30.3%      Mismatches: 27
Query Match: 14.4%      Indels: 17
DB: 9      Gaps: 3

US-09-917-376-4 (1-89) x US-11-072-512-279 (1-2564)
QY 23 ProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrVal 42
Db 454 CCCTCCGAAGCGATGGCTGCACACTCGCTCCAGCAGCTTGTCCCGCAGCTCTCGAATACC 395
QY 43 ArgTyrTTPheThr----- 47
Db 394 GACTGGTGGAACTCCAGCGCGCGTCCGTCGAGGTCCCAAGAGGGGCGTGA 335
QY 48 ArgAspGlyGlySerSerThrLeu---ValTyrAsnCysAspTTPAlaAlaMetGlyCys 66
Db 334 AGAGATGGCAGCAGCACACCATTTCTGTCTGTGAACTGCTCGATGCTTGGCGGCTCC 275
QY 67 GlyAsn---IleArgAlaSerPheGlySerValAsnProAlaThrPro 81
Db 274 GTGCAGTTGGTCAGGGTCTCTTTCAGGTCTCGCCGCTTGGCCACGCC 227

Search completed: March 2, 2006, 13:23:22
Job time : 103.506 secs
```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 13:55:49 ; Search time 5.0715 Seconds
(without alignments)
1434.576 Million cell updates/sec

Title: US-09-917-376-5
Perfect score: 469
Sequence: 1 VSGGVQVQKNDSPGNDQ.....IRASFGSVNPATPTADTYLQ 88

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5 COMB.pcp.*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pcp.*
3: /cgn2_6/ptodata/1/iaa/H COMB.pcp.*
4: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pcp.*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pcp.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	195.5	41.7	616	2	US-09-136-574A-47
2	194.5	41.5	1751	2	US-09-136-574A-44
3	191.5	40.8	1426	2	US-09-136-574A-43
4	172	36.7	700	1	US-07-862-588B-2
5	165.5	35.3	551	1	US-09-033-537A-1
6	165	35.2	167	4	PCT-US95-13813-9
7	165	35.2	476	2	US-09-339-159B-4
8	165	35.2	493	2	US-09-198-955A-10
9	165	35.2	493	2	US-09-198-955A-12
10	165	35.2	493	2	US-09-694-531-12
11	165	35.2	493	2	US-09-670-141-10
12	165	35.2	493	2	US-10-072-152-12
13	155.5	33.2	1352	2	US-09-784-554B-2
14	150.5	32.1	1350	2	US-09-784-554B-4
15	120.5	25.7	531	1	US-07-862-588B-7
16	114	24.3	162	1	US-08-048-164A-2
17	114	24.3	162	1	US-08-460-462-2
18	114	24.3	162	1	US-08-460-457-2
19	114	24.3	162	1	US-08-460-458-2
20	114	24.3	162	1	US-08-460-455-2
21	114	24.3	162	1	US-08-330-394A-2
22	114	24.3	163	2	US-09-006-636-7
23	114	24.3	163	2	US-09-006-632-7
24	114	24.3	163	2	US-09-325-274-7
25	113	24.1	382	2	US-09-277-716-22
26	113	24.1	382	2	US-09-609-161B-22
27	112	23.9	154	1	US-08-330-394A-29

28	112	23.9	156	1	US-08-330-394A-22	Sequence 22, Appl
29	72.5	15.5	272	2	US-09-902-540-11056	Sequence 11056, A
30	68.5	14.6	500	2	US-10-104-047-2825	Sequence 2825, Ap
31	65.5	14.0	1566	2	US-09-581-472B-2	Sequence 2, Appli
32	65	13.9	474	2	US-09-248-796A-25524	Sequence 25524, A
33	64	13.6	143	2	US-09-301-593-26	Sequence 26, Appl
34	64	13.6	428	2	US-09-118-319-5	Sequence 5, Appli
35	64	13.6	464	1	US-08-353-400-36	Sequence 36, Appl
36	64	13.6	472	2	US-09-301-593-30	Sequence 30, Appl
37	63.5	13.5	1785	2	US-09-341-587-3	Sequence 3, Appli
38	63	13.4	453	2	US-09-301-593-18	Sequence 18, Appl
39	63	13.4	472	2	US-09-301-593-43	Sequence 43, Appl
40	63	13.4	718	2	US-09-328-352-4640	Sequence 4640, Ap
41	63	13.4	1133	2	US-09-902-540-12243	Sequence 12243, A
42	63	13.4	1581	2	US-09-110-517-2	Sequence 2, Appli
43	62.5	13.3	288	2	US-09-423-439-38	Sequence 38, Appl
44	62.5	13.3	389	2	US-09-902-540-14347	Sequence 14347, A
45	62.5	13.3	445	1	US-08-353-400-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1

US-09-136-574A-47
; Sequence 47, Application US/09136574A
; Patent No. 6294366

GENERAL INFORMATION:

APPLICANT: Farrington, Graham K.
Anderson, Paige
Gibbs, Moreland
Bergquist, Peter
Daniels, Roy
Morgan, Hugh W.
Williams, Diane P.

TITLE OF INVENTION: Compositions and Methods for
Treating Cellulose Containing Fabrics Using Truncated
Cellulase Enzyme Compositions

NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Center, P.O. Box 457
CITY: Spring House
STATE: PA
COUNTRY: USA
ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/136,574A

FILING DATE: 19-Aug-1998

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/932,571

FILING DATE: September 19, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Bak, Mary E.

REGISTRATION NUMBER: 31,215

REFERENCE/DOCKET NUMBER: 1997US001/CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-540-9200

TELEFAX: 215-540-5818

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 47:

SEQUENCE CHARACTERISTICS:

LENGTH: 616 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 47:

US-09-136-574A-47

Query Match 41.7%; Score 195.5; DB 2; Length 616;
Best Local Similarity 43.2%; Pred. No. 3.4e-13;
Matches 38; Conservative 15; Mismatches 34; Indels 1; Gaps 1;

QY 1 VSGGVKVKYKNDSPAGNQIKPGLQVLNTGSSVDLSRVKIRYWTVDGKPSAV-CD 60
DB 1 MCGVKVLYKNNETSASTGSRPWFVKIYNGSSVDLSRVKIRYWTVDGKPSAV-CD 59
QY 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
DB 60 WAQIGASNVTFNFKVLSGSGVSGADYYLE 87

RESULT 2

US-09-136-574A-44
Sequence 44, Application US/09136574A
Patent No. 6294366

GENERAL INFORMATION:

APPLICANT: Farrington, Graham K.

Anderson, Paige

Gibbs, Moreland

Bergquist, Peter

Daniels, Roy

Morgan, Hugh W.

Williams, Diane P.

TITLE OF INVENTION: Compositions and Methods for

Treating Cellulose Containing Fabrics Using Truncated

Cellulase Enzyme Compositions

NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Center, P.O. Box 457

CITY: Spring House

STATE: PA

COUNTRY: USA

ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/136,574A

FILING DATE: 19-Aug-1998

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/932,571

FILING DATE: September 19, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Bak, Mary E.

REGISTRATION NUMBER: 31,215

REFERENCE/DOCKET NUMBER: 1997US001/CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-540-9200

TELEFAX: 215-540-5818

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 44:

SEQUENCE CHARACTERISTICS:

LENGTH: 1751 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 44:

US-09-136-574A-44

Query Match 41.5%; Score 194.5; DB 2; Length 1751;
Best Local Similarity 44.7%; Pred. No. 1.6e-12;
Matches 38; Conservative 14; Mismatches 32; Indels 1; Gaps 1;

QY 4 GVKVQYKNDSPAGNQIKPGLQVLNTGSSVDLSRVKIRYWTVDGKPSAV-CD 63
DB 4 GVKVQYKNDSPAGNQIKPGLQVLNTGSSVDLSRVKIRYWTVDGKPSAV-CD 63

DB 678 GVKVLYKNNETSASTGSRPWFVKIYNGSSVDLSRVKIRYWTVDGKPSAV-CD 736

QY 64 MCGNIRASFGSVNPATPTADTYLQ 88
DB 737 IGASNVTFNFKVLSGSGVSGADYYLE 761

RESULT 3

US-09-136-574A-43

Sequence 43, Application US/09136574A

Patent No. 6294366

GENERAL INFORMATION:

APPLICANT: Farrington, Graham K.

Anderson, Paige

Gibbs, Moreland

Bergquist, Peter

Daniels, Roy

Morgan, Hugh W.

Williams, Diane P.

TITLE OF INVENTION: Compositions and Methods for

Treating Cellulose Containing Fabrics Using Truncated

Cellulase Enzyme Compositions

NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Center, P.O. Box 457

CITY: Spring House

STATE: PA

COUNTRY: USA

ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/136,574A

FILING DATE: 19-Aug-1998

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/932,571

FILING DATE: September 19, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Bak, Mary E.

REGISTRATION NUMBER: 31,215

REFERENCE/DOCKET NUMBER: 1997US001/CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-540-9200

TELEFAX: 215-540-5818

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 43:

SEQUENCE CHARACTERISTICS:

LENGTH: 1426 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: No. 6294366e

SEQUENCE DESCRIPTION: SEQ ID NO: 43:

US-09-136-574A-43

Query Match 40.8%; Score 191.5; DB 2; Length 1426;
Best Local Similarity 43.5%; Pred. No. 2.6e-12;
Matches 37; Conservative 15; Mismatches 32; Indels 1; Gaps 1;

QY 4 GVKVQYKNDSPAGNQIKPGLQVLNTGSSVDLSRVKIRYWTVDGKPSAV-CD 63
DB 413 GLKVLKNNETSASTGSRPWFVKIYNGSSVDLSRVKIRYWTVDGKPSAV-CD 471

QY 64 MCGNIRASFGSVNPATPTADTYLQ 88
DB 472 IGASNVTFNFKVLSGSGVSGADYYLE 496

RESULT 4

```
US-07-862-588B-2
; Sequence 2, Application US/07862588B
; Patent No. 5916796
; GENERAL INFORMATION:
; APPLICANT: Joergensen, Per Linnaa
; APPLICANT: Sch lein, Martin
; APPLICANT: Hansen, Christian
; TITLE OF INVENTION: An Enzyme Exhibiting Cellulase Activity
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5916796o No. 5916796disk of No. 5916796th America, Inc.
; STREET: 405 Lexington Avenue, 62nd floor
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/862,588B
; FILING DATE: 19920727
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 164/90
; FILING DATE: 19-JAN-1990
; APPLICATION NUMBER: PCT/DK91/00013
; FILING DATE: 18-JAN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Zelson, Steve T. / Lambiris, Elias J.
; REGISTRATION NUMBER: 30,335 / 33,728
; REFERENCE/DOCKET NUMBER: 3425.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 700 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-862-588B-2

Query Match 36.7%; Score 172; DB 1; Length 700;
Best Local Similarity 44.9%; Pred. No. 1.5e-10;
Matches 40; Conservative 18; Mismatches 27; Indels 4; Gaps 3;

QY 1 VSGGVQVQKND-SAPGDNQIKPGLQVNTGSSVDLSTVTVRYWFTDGGSSTLVYNC 59
Db 549 VNSDLVQVQKGRNATDNQIKPHFNQKTSFVLSLTRYFTKD-SSAAMNGWI 607
QY 60 DWAAMCGNIRASFGSVNPATPTADTYLQ 88
Db 608 DWAUKGGSNIQISFGNHGA--DSDTYAE 634

RESULT 5
US-09-033-537A-1
; Sequence 1, Application US/09033537A
; Patent No. 5958083
; GENERAL INFORMATION:
; APPLICANT: Onishi, Masahiro
; APPLICANT: Fich, Merete
; APPLICANT: Toft, Annette Hanne
; APPLICANT: Sch lein, Martin
; TITLE OF INVENTION: Prevention Of Back-Staining
; TITLE OF INVENTION: In Stone Washing
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5958083o No. 5958083disk of No. 5958083th America, Inc.
```

```
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: U.S.A.
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/033,537A
FILING DATE: 02-MAR-1998
CLASSIFICATION: 008
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 0993/95
FILING DATE: 08-SEP-1995
APPLICATION NUMBER: PCT/DK96/00364
FILING DATE: 03-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4492.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 551 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-033-537A-1

Query Match 35.3%; Score 165.5; DB 1; Length 551;
Best Local Similarity 37.9%; Pred. No. 6e-10;
Matches 33; Conservative 22; Mismatches 29; Indels 3; Gaps 2;

QY 2 SGGVKVQKNDAPGDNQIKPGLQVNTGSSVDLSTVTVRYWFTDGGSSTLVYNC 61
Db 402 TGNLVQVQKVDTSATDNQMKPSFNKKNKTTFPVNLGKLRYYFTKD-GTADMSASPDW 460
QY 62 AAMCGNIRASFGSVNPATPTADTYLQ 88
Db 461 AQIGASNVSAAP--ANFTGSNTDTIYE 485

RESULT 6
PCT-US95-13813-9
; Sequence 9, Application PC/TUS9513813
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co. Ltd.
; APPLICANT: Ramot University Authority for Applied
; APPLICANT: Research and Industrial Development Ltd.
; APPLICANT: Technion Research and Development Foundation Ltd.
; APPLICANT: Bayer, Edward A.
; APPLICANT: Morag, Ely
; APPLICANT: Wilchek, Meir
; APPLICANT: Lamed, Raphael
; APPLICANT: Shoham, Yuval
; TITLE OF INVENTION: MODIFIED CELLULOSE-BINDING DOMAIN (CBD)
; TITLE OF INVENTION: PROTEINS AND USE THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brody and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; ; COMPUTER: IBM PC compatible
; ; OPERATING SYSTEM: PC-DOS/MS-DOS
; ; SOFTWARE: PatentIn Release #1.0, Version #1.30
; ; CURRENT APPLICATION DATA:
; ; APPLICATION NUMBER: PCT/US95/13813
; ; FILING DATE:
; ; CLASSIFICATION:
; ; ATTORNEY/AGENT INFORMATION:
; ; NAME: Browdy, Roger L.
; ; REGISTRATION NUMBER: 25,618
; ; REFERENCE/DOCKET NUMBER: BAYER=3
; ; TELECOMMUNICATION INFORMATION:
; ; TELEPHONE: (202) 628-5197
; ; TELEFAX: (202) 737-3528
; ; TELEX: 248633
; ; INFORMATION FOR SEQ ID NO: 9:
; ; SEQUENCE CHARACTERISTICS:
; ; LENGTH: 167 amino acids
; ; TYPE: amino acid
; ; STRANDEDNESS: single
; ; TOPOLOGY: linear
; ; MOLECULE TYPE: peptide
; ; PCT-US95-13813-9

Query Match 35.2%; Score 165; DB 4; Length 167;
Best Local Similarity 36.1%; Pred. No. 1.6e-10;
Matches 35; Conservative 23; Mismatches 29; Indels 10; Gaps 2;

QY 1 VSGGVKQVQKNDAPGDNQIKPGLQLVNTGSSVDLSTVTVRYWFTRDGSSSTLVNCD 60
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 5 VSGNLKVEFYNSPDDTNSINPQKVTNTGSSAIDLSKLTIRYYTVDGKQDTFW-CD 63
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 61 WAAM-----GCGNIRASFGSVNPTPTADTYLQ 88
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 64 HAAIGSGSYNGITSNVKGTFFVKMSSSTNNADTYLE 100
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 7
US-09-339-159B-4
; Sequence 4, Application US/09339159B
; Patent No. 6566114
; GENERAL INFORMATION:
; APPLICANT: Kauppinen, Markus
; APPLICANT: Schulein, Martin
; APPLICANT: Schnorr, Kirk
; APPLICANT: Andersen, Lene
; APPLICANT: Bjornvad, Mads
; TITLE OF INVENTION: No. 6566114el Mannanases
; FILE REFERENCE: 5440.204-US
; CURRENT APPLICATION NUMBER: US/09/339,159B
; CURRENT FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Bacillus
US-09-339-159B-4

Query Match 35.2%; Score 165; DB 2; Length 476;
Best Local Similarity 36.1%; Pred. No. 5.7e-10;
Matches 35; Conservative 23; Mismatches 29; Indels 10; Gaps 2;

QY 1 VSGGVKQVQKNDAPGDNQIKPGLQLVNTGSSVDLSTVTVRYWFTRDGSSSTLVNCD 60
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 314 VSGNLKVEFYNSPDDTNSINPQKVTNTGSSAIDLSKLTIRYYTVDGKQDTFW-CD 372
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 61 WAAM-----GCGNIRASFGSVNPTPTADTYLQ 88
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 373 HAAIGSGSYNGITSNVKGTFFVKMSSSTNNADTYLE 409
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 8
US-09-198-956-10
; Sequence 10, Application US/09198956
; Patent No. 6165769
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schulein, Martin
; APPLICANT: Lange, Niels Erik K.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Schnorr, Kirk
; TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus
; FILE REFERENCE: 5377.200-US
; CURRENT APPLICATION NUMBER: US/09/198,956
; CURRENT FILING DATE: 1998-11-24
; EARLIER APPLICATION NUMBER: 1344/97
; EARLIER FILING DATE: 1997-11-24
; EARLIER APPLICATION NUMBER: 60/067,240
; EARLIER FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-09-198-956-10

Query Match 35.2%; Score 165; DB 2; Length 493;
Best Local Similarity 36.1%; Pred. No. 5.9e-10;
Matches 35; Conservative 23; Mismatches 29; Indels 10; Gaps 2;

QY 1 VSGGVKQVQKNDAPGDNQIKPGLQLVNTGSSVDLSTVTVRYWFTRDGSSSTLVNCD 60
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 331 VSGNLKVEFYNSPDDTNSINPQKVTNTGSSAIDLSKLTIRYYTVDGKQDTFW-CD 389
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 61 WAAM-----GCGNIRASFGSVNPTPTADTYLQ 88
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 390 HAAIGSGSYNGITSNVKGTFFVKMSSSTNNADTYLE 426
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 9
US-09-198-955A-12
; Sequence 12, Application US/09198955A
; Patent No. 6187580
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schulein, Martin
; APPLICANT: Lange, Niels E.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Moller, Soren
; APPLICANT: Glad, Sanne O. S.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Schnorr, Kirk
; APPLICANT: Kongsbak, Lars
; TITLE OF INVENTION: No. 6187580el Pectate Lyases
; FILE REFERENCE: 5378.200-US
; CURRENT APPLICATION NUMBER: US/09/198,955A
; CURRENT FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: 1343/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 1344/97
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 09/073,684
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/067,249
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 60/067,240
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 09/073,684
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 09/184,217
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 493
```

```
; TYPE: PRT
; ORGANISM: Clostridium thermocellum
US-09-198-955A-12

Query Match      35.2%; Score 165; DB 2; Length 493;
Best Local Similarity 36.1%; Pred. No. 5.9e-10;
Matches 35; Conservative 23; Mismatches 29; Indels 10; Gaps 2;

Qy 1 VSGGVKQYKNDAPGDNQIKPGLQLVNTGSSVDLSTVTVRYWFTRDGSSSTLVYVNC 60
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 331 VSGNLKVEFYNSPDSITNSINPQKVTNTGSSAIDLSKLTLYYYTVDGQKQTFW-CD 389

Qy 61 WAAM-----CGCNIRASFGSVNPATPTADTYLQ 88
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 390 HAAIIGSNGSYNGITSNVKGTFFVKMSSTNNADTYLE 426

RESULT 10
US-09-694-531-12
; Sequence 12, Application US/09694531
; Patent No. 6368843
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schulein, Martin
; APPLICANT: Lange, Niels E.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Moller, Soren
; APPLICANT: Glad, Sanne O. S.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Schnorr, Kirk
; APPLICANT: Kongsbak, Lars
; TITLE OF INVENTION: No. 6368843el Pectate Lyases
; FILE REFERENCE: 5378.200-US
; CURRENT APPLICATION NUMBER: US/09/694,531
; CURRENT FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 09/198,955
; PRIOR FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: 1343/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 1344/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/067,249
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 60/067,240
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 09/073,684
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 09/184,217
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Clostridium thermocellum
US-09-694-531-12

Query Match      35.2%; Score 165; DB 2; Length 493;
Best Local Similarity 36.1%; Pred. No. 5.9e-10;
Matches 35; Conservative 23; Mismatches 29; Indels 10; Gaps 2;

Qy 1 VSGGVKQYKNDAPGDNQIKPGLQLVNTGSSVDLSTVTVRYWFTRDGSSSTLVYVNC 60
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 331 VSGNLKVEFYNSPDSITNSINPQKVTNTGSSAIDLSKLTLYYYTVDGQKQTFW-CD 389

Qy 61 WAAM-----CGCNIRASFGSVNPATPTADTYLQ 88
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 390 HAAIIGSNGSYNGITSNVKGTFFVKMSSTNNADTYLE 426

RESULT 11
US-09-670-141-10
; Sequence 10, Application US/09670141
```

```
; SEQ ID NO 12
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Clostridium thermocellum
US-10-072-152-12

Query Match      35.2%; Score 165; DB 2; Length 493;
Best Local Similarity 36.1%; Pred. No. 5.9e-10;
Matches 35; Conservative 23; Mismatches 29; Indels 10; Gaps 2;

Qy 1 VSGGVKQYKNDAPGDNQIKPGLQLVNTGSSVDLSTVTVRYWFTRDGSSSTLVNCD 60
Db 331 VSGNLKVEFNSPDSTNSINPQKVTNTSSAIDSKLTIRYYIVDQKQDTFW-CD 389

Qy 61 WAAM-----CGGNIRASFGSVNPATPTADTYLQ 88
Db 390 HAAIGNSGYNGITSNVKGTGFKVMSSTNNADIYLE 426

RESULT 13
US-09-784-554B-2
; Sequence 2, Application US/09784554B
; Patent No. 6815192
; GENERAL INFORMATION:
; APPLICANT: Schmorz, Kirk
; APPLICANT: Jorgensen, Per Lina
; APPLICANT: Schulein, Martin
; TITLE OF INVENTION: FAMILY 44 XYLOGLUCANASES
; FILE REFERENCE: 10017-200-US
; CURRENT APPLICATION NUMBER: US/09/784,554B
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1352
; TYPE: PRT
; ORGANISM: Paenibacillus polymyxa
US-09-784-554B-2

Query Match      33.2%; Score 155.5; DB 2; Length 1352;
Best Local Similarity 31.8%; Pred. No. 2.3e-08;
Matches 27; Conservative 21; Mismatches 36; Indels 1; Gaps 1;

Qy 4 GVKVQYKNDAPGDNQIKPGLQLVNTGSSVDLSTVTVRYWFTRDGSSSTLVNCDWAA 63
Db 1200 GLLLQYRTADTKVNDHNLNPFQIVNKGTTSPINELKIRYYIIDGDREQ-TFNCDYAT 1258

Qy 64 MCGGNIRASFGSVNPATPTADTYLQ 88
Db 1259 LSCSKLNGKLVKDKAATGADYYLE 1283

RESULT 14
US-09-784-554B-4
; Sequence 4, Application US/09784554B
; Patent No. 6815192
; GENERAL INFORMATION:
; APPLICANT: Schmorz, Kirk
; APPLICANT: Jorgensen, Per Lina
; APPLICANT: Schulein, Martin
; TITLE OF INVENTION: FAMILY 44 XYLOGLUCANASES
; FILE REFERENCE: 10017-200-US
; CURRENT APPLICATION NUMBER: US/09/784,554B
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1350
; TYPE: PRT
; ORGANISM: Paenibacillus polymyxa
US-09-784-554B-4

Query Match      32.1%; Score 150.5; DB 2; Length 1350;

Best Local Similarity 31.8%; Pred. No. 8.1e-08;
Matches 27; Conservative 21; Mismatches 36; Indels 1; Gaps 1;

Qy 4 GVKVQYKNDAPGDNQIKPGLQLVNTGSSVDLSTVTVRYWFTRDGSSSTLVNCDWAA 63
Db 1199 GLVLQYRTADTKVNDHNLNPFQILNKGTISVPINELKIRYYIIDGDREQ-TFNCDYAV 1257

Qy 64 MCGGNIRASFGSVNPATPTADTYLQ 88
Db 1258 LSCSKLNGKLVKDKAATGADYYLE 1282

RESULT 15
US-07-862-588B-7
; Sequence 7, Application US/07862588B
; Patent No. 5916796
; GENERAL INFORMATION:
; APPLICANT: Joergensen, Per Linaa
; APPLICANT: Schlein, Martin
; APPLICANT: Hansen, Christian
; TITLE OF INVENTION: An Enzyme Exhibiting Cellulase Activity
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5916796o No. 5916796disk of No. 5916796th America, Inc.
; STREET: 405 Lexington Avenue, 62nd floor
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/862,588B
; FILING DATE: 19920727
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 164/90
; FILING DATE: 19-JAN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/DK91/00013
; FILING DATE: 18-JAN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Zelson, Steve T. / Lambiris, Elias J.
; REGISTRATION NUMBER: 30,335 / 33,728
; REFERENCE/DOCKET NUMBER: 3425.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 531 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-862-588B-7

Query Match      25.7%; Score 120.5; DB 1; Length 531;
Best Local Similarity 32.9%; Pred. No. 5.2e-05;
Matches 28; Conservative 18; Mismatches 36; Indels 3; Gaps 2;

Qy 2 SGGVKQYKNDAPGDNQIKPGLQLVNTGSSVDLSTVTVRYWFTRDGSSSTLVNCDW 61
Db 432 TGNLVQYKVGDTSATDNQMKPSFNKKNKTPVNLGLKXXXXXKD-GPADMSCSIDW 490

Qy 62 AAMCGGNIRASFGSVNPATPTADTY 86
Db 491 AQIGRTNVLLAF--ANFTGSNTDTY 513

Search completed: March 2, 2006, 13:58:00
```

Job time : 5.0715 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 14:20:38 ; Search time 16.952 Seconds
(without alignments)
2169.009 Million cell updates/sec

Title: US-09-917-376-5
Perfect score: 469
Sequence: 1 VSGGVKQVKNDSAPGDNQ.....TRASFGSVNPATPTADTYLQ 88

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pap.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pap.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pap.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pap.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pap.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	469	100.0	88	3	US-09-917-376-5
2	469	100.0	88	4	US-10-155-400-5
3	469	100.0	89	3	US-09-917-376-4
4	469	100.0	89	4	US-10-155-400-4
5	469	100.0	154	3	US-09-917-378-4
6	469	100.0	762	3	US-09-917-378-1
7	466	99.4	150	3	US-09-917-384-5
8	466	99.4	150	3	US-09-917-383-5
9	466	99.4	1043	3	US-09-917-383-6
10	466	99.4	1043	3	US-09-917-383-6
11	466	99.4	1228	3	US-09-917-384-1
12	466	99.4	1228	3	US-09-917-383-1
13	462	98.5	957	3	US-09-917-376-1
14	462	98.5	957	4	US-10-155-400-1
15	257	54.8	741	4	US-10-156-761-8100
16	190	40.5	508	4	US-10-369-493-23151
17	188	40.1	477	4	US-10-466-208-12
18	188	40.1	496	4	US-10-466-208-8
19	188	40.1	677	4	US-10-433-577-35
20	180	38.4	1621	4	US-10-185-990-10
21	165	35.2	167	5	US-10-933-404-4
22	165	35.2	476	4	US-10-372-054-4
23	165	35.2	493	4	US-10-072-152-12
24	165	35.2	493	4	US-10-653-433-12
25	165	35.2	599	3	US-09-955-555A-29
26	155.5	33.2	1352	3	US-09-784-5548-2
27	155.5	33.2	1352	5	US-10-896-555-2

28	150.5	32.1	1350	3	US-09-784-554B-4	Sequence 4, Appli
29	150.5	32.1	1350	5	US-10-896-555-4	Sequence 4, Appli
30	130.5	27.8	1483	4	US-10-282-122A-51483	Sequence 51483, A
31	114	24.3	256	4	US-10-261-446-6	Sequence 6, Appli
32	114	24.3	256	4	US-10-261-445B-6	Sequence 6, Appli
33	114	24.3	256	5	US-10-782-234-6	Sequence 6, Appli
34	113	24.1	382	3	US-09-808-898-22	Sequence 22, Appli
35	105.5	22.5	163	4	US-10-460-524-2	Sequence 2, Appli
36	78.5	16.7	1049	4	US-10-282-122A-49900	Sequence 49900, A
37	75.5	16.1	346	4	US-10-437-963-179234	Sequence 179234,
38	75.5	16.1	618	4	US-10-211-462-223	Sequence 223, App
39	75.5	16.1	618	5	US-10-723-860-4039	Sequence 4039, Ap
40	75.5	16.1	618	5	US-10-756-149-5713	Sequence 5713, Ap
41	75.5	16.1	986	3	US-09-747-835A-53	Sequence 53, Appl
42	75.5	16.1	986	4	US-10-120-604-101	Sequence 101, App
43	75.5	16.1	986	4	US-10-325-567A-406	Sequence 406, App
44	75.5	16.1	986	4	US-10-398-458-3	Sequence 3, Appli
45	75.5	16.1	986	4	US-10-312-312-53	Sequence 53, Appl

ALIGNMENTS

RESULT 1

US-09-917-376-5
; Sequence 5, Application US/09917376
; Publication No. US20040038334A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOLYTICUS
; FILE REFERENCE: 40197.4US01
; CURRENT APPLICATION NUMBER: US/09/917,376
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
; FEATURE:
; OTHER INFORMATION: Carbohydrate binding domain
US-09-917-376-5

Query Match		100.0%;	Score 469;	DB 3;	Length 88;
Best Local Similarity		100.0%;	Pred. No. 3.6e-48;		
Matches		88;	Conservative	0;	Mismatches
				0;	Indels
				0;	Gaps
				0;	
Qy	1	VSGGVKQVKNDSAPGDNQIKPGLQVNTGSSVDLSTVTYVWFRDGSSTLVNCD	60		
			</		

RESULT 2

US-10-155-400-5
; Sequence 5, Application US/10155400
; Publication No. US2003010898A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOLYTICUS
; FILE REFERENCE: NREL 01-36A
; CURRENT APPLICATION NUMBER: US/10/155,400

; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
; FEATURE:
; OTHER INFORMATION: Carbohydrate binding domain
US-10-155-400-5

Query Match 100.0%; Score 469; DB 4; Length 88;
Best Local Similarity 100.0%; Pred. No. 3.6e-48;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VSGGVKQVQKNDAPGDNQIKPGLQLVNTGSSVDLSTVTVRYWFTDRDGSSTLVNCD 60
DB 1 VSGGVKQVQKNDAPGDNQIKPGLQLVNTGSSVDLSTVTVRYWFTDRDGSSTLVNCD 60
QY 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
DB 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88

RESULT 3
US-09-917-376-4
; Sequence 4, Application US/09917376
; Publication No. US20040038334A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOYTICUS
; FILE REFERENCE: 40197.4US01
; CURRENT APPLICATION NUMBER: US/09/917.376
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (89)
; OTHER INFORMATION: Any amino acid
; OTHER INFORMATION: Carbohydrate binding domain
US-09-917-376-4

Query Match 100.0%; Score 469; DB 3; Length 89;
Best Local Similarity 100.0%; Pred. No. 3.7e-48;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VSGGVKQVQKNDAPGDNQIKPGLQLVNTGSSVDLSTVTVRYWFTDRDGSSTLVNCD 60
DB 1 VSGGVKQVQKNDAPGDNQIKPGLQLVNTGSSVDLSTVTVRYWFTDRDGSSTLVNCD 60
QY 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
DB 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88

RESULT 4
US-10-155-400-4
; Sequence 4, Application US/10155400
; Publication No. US20030108988A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.

; TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOYTICUS
; FILE REFERENCE: NREL 01-36A
; CURRENT APPLICATION NUMBER: US/10/155,400
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (89)
; OTHER INFORMATION: Any amino acid
; FEATURE:
; OTHER INFORMATION: Carbohydrate binding domain
US-10-155-400-4

Query Match 100.0%; Score 469; DB 4; Length 89;
Best Local Similarity 100.0%; Pred. No. 3.7e-48;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VSGGVKQVQKNDAPGDNQIKPGLQLVNTGSSVDLSTVTVRYWFTDRDGSSTLVNCD 60
DB 1 VSGGVKQVQKNDAPGDNQIKPGLQLVNTGSSVDLSTVTVRYWFTDRDGSSTLVNCD 60
QY 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
DB 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88

RESULT 5
US-09-917-378-4
; Sequence 4, Application US/09917378
; Publication No. US20030119093A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT MANNANASE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOYTICUS
; FILE REFERENCE: 40197.7US01
; CURRENT APPLICATION NUMBER: US/09/917,378
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Carbohydrate
; OTHER INFORMATION: binding domain
US-09-917-378-4

Query Match 100.0%; Score 469; DB 3; Length 154;
Best Local Similarity 100.0%; Pred. No. 7.1e-48;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VSGGVKQVQKNDAPGDNQIKPGLQLVNTGSSVDLSTVTVRYWFTDRDGSSTLVNCD 60
DB 1 VSGGVKQVQKNDAPGDNQIKPGLQLVNTGSSVDLSTVTVRYWFTDRDGSSTLVNCD 60
QY 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
DB 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
RESULT 6
US-09-917-378-1

Sequence 1, Application US/09917378
 Publication No. US2003011903A1
 GENERAL INFORMATION:
 APPLICANT: DING, SHI-YOU
 APPLICANT: ADNEY, WILLIAM S.
 APPLICANT: VINZANT, TODD B.
 APPLICANT: DECKER, STEPHEN R.
 APPLICANT: HIMMEL, MICHAEL E.
 TITLE OF INVENTION: THERMAL TOLERANT MANNANASE FROM ACIDOTHERMUS
 TITLE OF INVENTION: CELLULOXYTICUS
 FILE REFERENCE: 40197.7US01
 CURRENT APPLICATION NUMBER: US/09/917,378
 CURRENT FILING DATE: 2001-07-28
 NUMBER OF SEQ ID NOS: 8
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 1
 LENGTH: 762
 TYPE: PRT
 ORGANISM: Acidothermus cellulolyticus
 US-09-917-378-1

Query Match 100.0%; Score 469; DB 3; Length 762;
 Best Local Similarity 100.0%; Pred. No. 4.9e-47;
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VSGGVKQVYKNNDSAPGDNQIKPGLQLVNTGSSVDLSVTYVRYWFTTRDGGSSTLVYNC 60
 DB 455 VSGGVKQVYKNNDSAPGDNQIKPGLQLVNTGSSVDLSVTYVRYWFTTRDGGSSTLVYNC 514
 QY 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
 DB 515 WAAMCGNIRASFGSVNPATPTADTYLQ 542

RESULT 7
 US-09-917-384-5
 Sequence 5, Application US/09917384
 Publication No. US20030096342A1
 GENERAL INFORMATION:
 APPLICANT: DING, SHI-YOU
 APPLICANT: ADNEY, WILLIAM S.
 APPLICANT: VINZANT, TODD B.
 APPLICANT: DECKER, STEPHEN R.
 APPLICANT: HIMMEL, MICHAEL E.
 TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE FROM ACIDOTHERMUS
 TITLE OF INVENTION: CELLULOXYTICUS
 FILE REFERENCE: 40170.6US01
 CURRENT APPLICATION NUMBER: US/09/917,384
 CURRENT FILING DATE: 2001-07-28
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 5
 LENGTH: 150
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: Segment of
 OTHER INFORMATION: GUXA
 US-09-917-384-5

Query Match 99.4%; Score 466; DB 3; Length 150;
 Best Local Similarity 98.9%; Pred. No. 1.6e-47;
 Matches 87; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VSGGVKQVYKNNDSAPGDNQIKPGLQLVNTGSSVDLSVTYVRYWFTTRDGGSSTLVYNC 60
 DB 1 VSGGLKVQYKNNDSAPGDNQIKPGLQLVNTGSSVDLSVTYVRYWFTTRDGGSSTLVYNC 60
 QY 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
 DB 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88

RESULT 8
 US-09-917-383-5
 Sequence 5, Application US/09917383
 Publication No. US20030104522A1
 GENERAL INFORMATION:
 APPLICANT: DING, SHI-YOU
 APPLICANT: ADNEY, WILLIAM S.
 APPLICANT: VINZANT, TODD B.
 APPLICANT: DECKER, STEPHEN R.
 APPLICANT: HIMMEL, MICHAEL E.
 TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE FROM ACIDOTHERMUS
 TITLE OF INVENTION: CELLULOXYTICUS
 FILE REFERENCE: 40170.6US01
 CURRENT APPLICATION NUMBER: US/09/917,383
 CURRENT FILING DATE: 2001-07-28
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 5
 LENGTH: 150
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: Segment of
 OTHER INFORMATION: GUXA
 US-09-917-383-5

Query Match 99.4%; Score 466; DB 3; Length 150;
 Best Local Similarity 98.9%; Pred. No. 1.6e-47;
 Matches 87; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VSGGVKQVYKNNDSAPGDNQIKPGLQLVNTGSSVDLSVTYVRYWFTTRDGGSSTLVYNC 60
 DB 1 VSGGLKVQYKNNDSAPGDNQIKPGLQLVNTGSSVDLSVTYVRYWFTTRDGGSSTLVYNC 60
 QY 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
 DB 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88

RESULT 9
 US-09-917-384-6
 Sequence 6, Application US/09917384
 Publication No. US20030096342A1
 GENERAL INFORMATION:
 APPLICANT: DING, SHI-YOU
 APPLICANT: ADNEY, WILLIAM S.
 APPLICANT: VINZANT, TODD B.
 APPLICANT: DECKER, STEPHEN R.
 APPLICANT: HIMMEL, MICHAEL E.
 TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE FROM ACIDOTHERMUS
 TITLE OF INVENTION: CELLULOXYTICUS
 FILE REFERENCE: 40170.6US01
 CURRENT APPLICATION NUMBER: US/09/917,384
 CURRENT FILING DATE: 2001-07-28
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 6
 LENGTH: 1043
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: Segment of
 OTHER INFORMATION: GUXA
 US-09-917-384-6

Query Match 99.4%; Score 466; DB 3; Length 1043;
 Best Local Similarity 98.9%; Pred. No. 1.7e-46;
 Matches 87; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VSGGVKQVYKNNDSAPGDNQIKPGLQLVNTGSSVDLSVTYVRYWFTTRDGGSSTLVYNC 60
 DB 477 VSGGLKVQYKNNDSAPGDNQIKPGLQLVNTGSSVDLSVTYVRYWFTTRDGGSSTLVYNC 536

Qy	61	WAAMGCGNIRASFGSVNPATPTADTYLQ 88
Db	537	WAAMGCGNIRASFGSVNPATPTADTYLQ 564

RESULT 10

US-09-917-383-6
; Sequence 6, Application US/09917383
; Publication NO. US20030104522A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE
; TITLE OF INVENTION: CELLULOSE TOLERANT CELLULOSE
; FILE REFERENCE: 40170.6US01
; CURRENT APPLICATION NUMBER: US/09/917,383
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Segment of
; OTHER INFORMATION: Guxa
US-09-917-383-6

RECIT.T 11

```

RES001 11
US-09-917-384-1
; Sequence 1, Application US/09917384
; Publication NO. US20030096342A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOLYTICUS
; FILE REFERENCE: 40170.6US01
; CURRENT APPLICATION NUMBER: US/09/917,384
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1228
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Segment of
; OTHER INFORMATION: Guxa
US-09-917-384-1

```

RESULT 13

```

US-09-917-376-1
; Sequence 1, Application US/09917376
; Publication No. US20040038334A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOLYTICUS
; FILE REFERENCE: 40197.4US01
; CURRENT APPLICATION NUMBER: US/09/917,376
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 957
; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (957)
; OTHER INFORMATION: Any amino acid

```

US-09-917-376-1

Query Match 98.5%; Score 462; DB 3; Length 957;
Best Local Similarity 97.7%; Pred. No. 4.5e-46;
Matches 86; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 VSGGVKQYKNDSPAGDNOIKPGLQVNTGSSVDLSTVTYRYWFTDGGSSSTLVYNC 60
DB 869 VSGGVKQYKNDSPAGDNOIKPGLQVNTGSSVDLSTVTYRYWFTDGGSSSTLVYNC 928
QY 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
DB 929 WAAIGCGNIRASFGSVNPATPTADTYLQ 956

RESULT 14

US-10-155-400-1
; Sequence 1, Application US/10155400
; Publication No. US20030108988A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOGLYTICUS
; FILE REFERENCE: NREL 01-36A
; CURRENT APPLICATION NUMBER: US/10/155,400
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 957
; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (957)
; OTHER INFORMATION: Any amino acid
US-10-155-400-1

Query Match 98.5%; Score 462; DB 4; Length 957;
Best Local Similarity 97.7%; Pred. No. 4.5e-46;
Matches 86; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 VSGGVKQYKNDSPAGDNOIKPGLQVNTGSSVDLSTVTYRYWFTDGGSSSTLVYNC 60
DB 869 VSGGVKQYKNDSPAGDNOIKPGLQVNTGSSVDLSTVTYRYWFTDGGSSSTLVYNC 928
QY 61 WAAMCGNIRASFGSVNPATPTADTYLQ 88
DB 929 WAAIGCGNIRASFGSVNPATPTADTYLQ 956

RESULT 15

US-10-156-761-8100
; Sequence 8100, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697

; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8100
; LENGTH: 741
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8100

Query Match 54.8%; Score 257; DB 4; Length 741;
Best Local Similarity 56.3%; Pred. No. 1.2e-21;
Matches 49; Conservative 13; Mismatches 25; Indels 0; Gaps 0;
QY 2 SGGVKVQYKNDSPAGDNOIKPGLQVNTGSSVDLSTVTYRYWFTDGGSSSTLVYNC 61
DB 591 SGGKLVLYKNDSSATDNAIRPGLRIVNTGSGSLDLSKVTRYFTTRDSSPTVNAWCDY 650
QY 62 AAMGCGNIRASFGSVNPATPTADTYLQ 88
DB 651 AAVGCSNVSLKVVPLTTPVPGADAYLE 677

Search completed: March 2, 2006, 14:26:55
Job time: 16.952 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 14:27:20 ; Search time 1.78442 Seconds
(without alignments)
986.322 Million cell updates/sec

Title: US-09-917-376-5

Perfect score: 469

Sequence: 1 VSGGVKVOYKNDSPAGDNQ.....IRASFGSVNPTPTADTYLQ 88

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 135339 seqs, 20000136 residues

Total number of hits satisfying chosen parameters: 135339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA_New.*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pdb.*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pdb.*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pdb.*
4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pdb.*
5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pdb.*
6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pdb.*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pdb.*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pdb.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	163	34.8	499	6	US-10-517-939-328
2	157	33.5	534	6	US-10-510-386-230
3	113	24.1	382	7	US-11-179-411-22
4	113	24.1	382	7	US-11-175-766-22
5	108	23.0	738	6	US-10-517-939-344
6	72	15.4	210	7	US-11-170-653-51
7	70	14.9	5291	7	US-11-052-554A-281
8	68.5	14.6	500	7	US-11-072-512-2825
9	68	14.5	755	6	US-10-517-933-330
10	64.5	13.8	389	7	US-11-207-626A-26
11	62	13.2	389	7	US-11-207-626A-20
12	61	13.0	350	6	US-10-517-939-14
13	61	13.0	1889	7	US-11-102-476-46
14	60.5	12.9	806	7	US-11-098-686-10278
15	60.5	12.9	983	7	US-11-087-099-7483
16	60	12.8	86	7	US-11-102-476-24
17	60	12.8	91	7	US-11-102-476-9
18	60	12.8	919	7	US-11-074-176-284
19	60	12.8	1225	7	US-11-102-476-2
20	60	12.8	1424	7	US-11-102-476-4
21	59.5	12.7	217	6	US-10-454-437-82
22	59	12.6	111	7	US-11-072-512-2820
23	59	12.6	124	6	US-10-771-257-19
24	59	12.6	124	7	US-11-127-677-19
25	59	12.6	182	7	US-11-170-653-49

26 59 12.6 208 7 US-11-214-413-32 Sequence 32, Appl
27 59 12.6 381 7 US-11-087-099-3485 Sequence 3485, Ap
28 59 12.6 498 7 US-11-131-479-76 Sequence 76, Appl
29 58.5 12.5 211 7 US-11-170-653-50 Sequence 50, Appl
30 58.5 12.5 228 7 US-11-170-653-39 Sequence 39, Appl
31 58.5 12.5 248 7 US-11-054-515-1790 Sequence 1790, Ap
32 58.5 12.5 327 7 US-11-064-246-15 Sequence 15, Appl
33 58.5 12.5 327 7 US-11-064-246-16 Sequence 16, Appl
34 58.5 12.5 345 7 US-11-107-028-37 Sequence 37, Appl
35 58 12.4 121 7 US-11-107-028-35 Sequence 35, Appl
36 58 12.4 130 7 US-11-188-187A-3 Sequence 3, Appl
37 58 12.4 140 7 US-11-183-205-62 Sequence 62, Appl
38 58 12.4 378 7 US-11-087-099-8872 Sequence 8872, Ap
39 58 12.4 378 7 US-11-087-099-11183 Sequence 1183, A
40 58 12.4 391 7 US-11-207-626A-32 Sequence 32, Appl
41 58 12.4 451 7 US-11-124-620-5 Sequence 5, Appl
42 58 12.4 451 7 US-11-124-620-7 Sequence 7, Appl
43 58 12.4 464 7 US-11-087-099-7240 Sequence 7240, Ap
44 58 12.4 2204 7 US-11-052-554A-134 Sequence 134, App
45 57.5 12.3 108 7 US-11-097-812-39 Sequence 39, Appl

ALIGNMENTS

RESULT 1
US-10-517-939-328
; Sequence 328, Application US/10517939
; Publication No. US20060003433A1
; GENERAL INFORMATION:
; APPLICANT: Steer, Brian
; APPLICANT: Callen, Walter
; APPLICANT: Healey, Shaun
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Wu, Di
; APPLICANT: Blum, David
; APPLICANT: Esteghalian, Alireza
; TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM
; FILE REFERENCE: 564462007901
; CURRENT APPLICATION NUMBER: US/10/517,939
; CURRENT FILING DATE: 2004-12-13
; PRIOR APPLICATION NUMBER: PCT/US03/19153
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/389,299
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 380
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 328
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample.
; NAME/KEY: SIGNAL
; LOCATION: (1)...(29)
US-10-517-939-328

Query Match 34.8%; Score 163; DB 6; Length 499;
Best Local Similarity 35.3%; Pred. No. 9.3e-12;
Matches 30; Conservative 21; Mismatches 30; Indels 4; Gaps 2;
QY 5 VKVQYKNDSPAGDNQIKPGLQVNTGGSSVDISTVTYVWF-TRDGGSTLVYNCWAA 63
DB 354 LSVQYRTGDSVNSNQIRPQIHVKNSKTNLKNLKVTVRYWYNTKNGQN---PDCDYAK 410
QY 64 MCGGNIRASFGSVNPTPTADTYLQ 88
DB 411 ICGSNVTHKFTVLQKPKVGADAYLE 435
RESULT 2

; FILE REFERENCE: 564462007901
; CURRENT APPLICATION NUMBER: US/10/517,939
; CURRENT FILING DATE: 2004-12-13
; PRIOR APPLICATION NUMBER: PCT/US03/19153
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/389,299
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 380
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 344
; LENGTH: 738
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample.
US-10-517-939-344

Query Match 23.0%; Score 108; DB 6; Length 738;
Best Local Similarity 32.1%; Pred. No. 5.2e-05;
Matches 27; Conservative 18; Mismatches 31; Indels 8; Gaps 4;

QY 7 VOYKNDAPGDN--QIKPGLQLVNTGSSVDLSTVTYRYWFTRDGSGSTLVYNCDAAM 64
DB 591 LQYR--SADGNNYQMKPQTIKNAGKVPILSELTIYYFTPE-STQPVDTRIDNAQF 646

QY 65 GCGNIRASFGSVNPTPTADTYLQ 88
DB 647 GAHVQTTV--VPPSDDAAAHAYVE 668

RESULT 6

US-11-170-653-51
; Sequence 51, Application US/11170653
; Publication No. US20050271769A1
; GENERAL INFORMATION:

; APPLICANT: Danisco A/S
; APPLICANT: Sibbesen, Ole
; APPLICANT: Sorensen, Jens
; TITLE OF INVENTION: Xylanase Variants Having Altered Sensitivity to Xylanase Inhibitor
; FILE REFERENCE: 674509-2046
; CURRENT APPLICATION NUMBER: US/11/170,653
; CURRENT FILING DATE: 2005-06-23
; PRIOR APPLICATION NUMBER: US/10/237,386
; PRIOR FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: PCT/IB01/00426
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: GB 0005585.5
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: GB 0015751.1
; PRIOR FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 51
; LENGTH: 210
; TYPE: PRT
; ORGANISM: A. tubigenis
US-11-170-653-51

Query Match 15.4%; Score 72; DB 7; Length 210;
Best Local Similarity 31.5%; Pred. No. 0.23;
Matches 23; Conservative 10; Mismatches 22; Indels 18; Gaps 4;

QY 29 NTGSSVDLSTVTYRYWFTRDG-----GSSLVYNCDAAMCGNIRASFG 74
DB 40 NLGDTYDESAGTFSMYW--EDGVSDFFVWGLWGTGSSSTITYSAEYASGSASLYAVYG 97

QY 75 SVNPTPTADTYL 87
DB 98 WVN--YPOAEYII 108

RESULT 7

US-11-052-554A-281

; Sequence 281, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:

; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 281
; LENGTH: 5291
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-281

Query Match 14.9%; Score 70; DB 7; Length 5291;
Best Local Similarity 33.3%; Pred. No. 18;
Matches 29; Conservative 15; Mismatches 19; Indels 24; Gaps 7;

QY 17 GD---NQIKPGLQLVNTGSS--VDL-STVTYRYWFTRDG---SSTLVYNCDA----- 61
DB 848 GDDVINAVERGETLVVSGTSGVEAGQTVTTF-----GSKNTTIVEANGSWTVNVPPA 902

QY 62 ---AAM--GCGNIRASFGSVNPTPTAD 84
DB 903 DLAALPDGAGNVQASVSNINGNSAQAD 929

RESULT 8

US-11-072-512-2825
; Sequence 2825, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2825
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-2825

```
Query Match          14.6%; Score 68.5; DB 7; Length 500;
Best Local Similarity 33.3%; Pred. No. 1.7;
Matches 19; Conservative 9; Mismatches 18; Indels 11; Gaps 3;

Qy 31 GSSGVDLSTVTVRYWFTRDG-----GSS--TLVYNCD----WAAMGCNIRASFGSV 76
Db 414 GTPETDIDSSCRYLKADGTQCPGSGSGTIVITCFEISAYGARGSANIKVTFISV 470

RESULT 9
US-10-517-939-330
; Sequence 330, Application US/10517939
; Publication No. US20060003433A1
; GENERAL INFORMATION:
; APPLICANT: Steer, Brian
; APPLICANT: Callen, Walter
; APPLICANT: Healey, Shaun
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Wu, Di
; APPLICANT: Blum, David
; APPLICANT: Esteghlalian, Alireza
; TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM
; FILE REFERENCE: 564462007901
; CURRENT APPLICATION NUMBER: US/10/517,939
; PRIOR FILING DATE: 2004-12-13
; PRIOR APPLICATION NUMBER: PCT/US03/19153
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/389,299
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 380
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 330
; LENGTH: 755
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample.
; NAME/KEY: SIGNAL
; LOCATION: (1)...(35)
US-10-517-939-330

Query Match          14.5%; Score 68; DB 6; Length 755;
Best Local Similarity 30.1%; Pred. No. 3.1;
Matches 31; Conservative 8; Mismatches 40; Indels 24; Gaps 5;

Qy 2 SGGVKVOYK-NNDAPGDNOIKPGLQVLNTGSSVDLSTVTVRYWFTRD----- 49
Db 651 SGACTVAYAITNDWGS-----FTANVTLTNTGSA--LNGWTLAYAFPGNOTISNAWNGTA 705

Qy 50 --GSSSTLVYNCDAWAMGCNIRASFG-----SVNPATPTADT 85
Db 706 VQSGSSVSVTNAGMNGSLPPNVASFGFQASYSNGNSVPSAFT 748

RESULT 10
US-11-207-626A-26
; Sequence 26, Application US/11207626A
; Publication No. US20060014276A1
; GENERAL INFORMATION:
; APPLICANT: Havenga, Menzo
; APPLICANT: Vogels, Ronald
; APPLICANT: Bout, Abraham
; TITLE OF INVENTION: CHIMERIC ADENOVIRUSES
; FILE REFERENCE: 2578-4123.2US
; CURRENT APPLICATION NUMBER: US/11/207,626A
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: EP 98202297.2
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26
```

```
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Human Adenovirus 32 Fiber Protein
US-11-207-626A-26

Query Match          13.8%; Score 64.5; DB 7; Length 389;
Best Local Similarity 32.0%; Pred. No. 3.8;
Matches 24; Conservative 10; Mismatches 32; Indels 9; Gaps 3;

Qy 11 NNDAPGDNOIKPGLQVLNTGSSVDLSTVTVRYWFTRDGSSSTLVYNCDAWAMGCNIR 70
Db 251 NNNTNPTDKKITVKL-LFNEKGVLMDSLSLKEYWNYRNDNSTSQAY--DNAVPFMPNLIK 307

Qy 71 ASFGSVNPATPTADT 85
Db 308 AY-----PKPTTDT 316

RESULT 11
US-11-207-626A-20
; Sequence 20, Application US/11207626A
; Publication No. US20060014276A1
; GENERAL INFORMATION:
; APPLICANT: Havenga, Menzo
; APPLICANT: Vogels, Ronald
; APPLICANT: Bout, Abraham
; TITLE OF INVENTION: CHIMERIC ADENOVIRUSES
; FILE REFERENCE: 2578-4123.2US
; CURRENT APPLICATION NUMBER: US/11/207,626A
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: EP 98202297.2
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Human Adenovirus 24 Fiber Protein
US-11-207-626A-20

Query Match          13.2%; Score 62; DB 7; Length 389;
Best Local Similarity 33.3%; Pred. No. 7.5;
Matches 25; Conservative 9; Mismatches 33; Indels 8; Gaps 3;

Qy 11 NNDAPGDNOIKPGLQVLNTGSSVDLSTVTVRYWFTRDGSSSTLVYNCDAWAMGCNIR 70
Db 250 NNNTNPTDKKITVKL-LFNEKGVLMDSLSLKEYWNYRNDNSTSQAY--DNAVPFMPNLIK 307

Qy 71 ASFGSVNPATPTADT 85
Db 308 AY-----PKPTTDT 316

RESULT 12
US-10-517-939-14
; Sequence 14, Application US/10517939
; Publication No. US20060003433A1
; GENERAL INFORMATION:
; APPLICANT: Steer, Brian
; APPLICANT: Callen, Walter
; APPLICANT: Healey, Shaun
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Wu, Di
; APPLICANT: Blum, David
; APPLICANT: Esteghlalian, Alireza
; TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM
; FILE REFERENCE: 564462007901
; CURRENT APPLICATION NUMBER: US/10/517,939
; CURRENT FILING DATE: 2004-12-13
; PRIOR APPLICATION NUMBER: PCT/US03/19153
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/389,299
```



```
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 380
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
US-10-917-939-14

Query Match      13.0%; Score 61; DB 6; Length 350;
Best Local Similarity 26.7%; Pred. No. 8.7;
Matches 16; Conservative 10; Mismatches 26; Indels 8; Gaps 1;

Qy 10 KNDSAPGDNQIKPGLQVNTGSSVDLSTVTVRVYFTRDGGSSSTLVYNCOW 61
Db 40 RHNSLTADNQMPESVLDRTATLAKGDLHAADVTRVDALMYFARDNGIPWRYHTLAW 99

RESULT 13
US-11-102-476-46
; Sequence 46, Application US/11102476
; Publication No. US2005021680A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Palaniappan, Raghavan U.M.
; APPLICANT: Cornell Research Foundation, Inc.
; TITLE OF INVENTION: Novel Immunogenic Proteins of Leptospira
; FILE REFERENCE: 1153.080US1
; CURRENT APPLICATION NUMBER: US/11/102,476
; CURRENT FILING DATE: 2005-04-08
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: PCT/US2003/32385
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/417721
; PRIOR FILING DATE: 2002-10-10
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 1889
; TYPE: PRT
; ORGANISM: Leptospira interrogans
US-11-102-476-46

Query Match      13.0%; Score 61; DB 7; Length 1889;
Best Local Similarity 26.2%; Pred. No. 62;
Matches 21; Conservative 14; Mismatches 27; Indels 18; Gaps 4;

Qy 19 NQIKPGL-----QLVNTG-----SSVDLSTVTVRVYFTRDGGSSSTLVY-----NCDWA 62
Db 498 NPVNPSLAKGLTKFTATGIYSNKNKDIITSAVT--WFSSDSSIIATISNAQKNQCNAYGA 555

Qy 63 AMCGNIRASFGSVNPATPT 82
Db 556 ATGTTDIKATFGKVSPPVST 575

RESULT 14
US-11-098-686-10278
; Sequence 10278, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433

; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10278
; LENGTH: 806
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10278

Query Match      12.9%; Score 60.5; DB 7; Length 806;
Best Local Similarity 33.3%; Pred. No. 26;
Matches 20; Conservative 5; Mismatches 20; Indels 15; Gaps 4;

Qy 10 KNDSAPGDNQIKPGLQVNTGSSVDLSTVTVRVYF--TRDQ--GSSSTLVY--NCDWAA 63
Db 286 KNNIVTPGKDPYPSPSSQLVRNNHQ-----RYWVRGMADGFIGNSTWCIISNLDYVS 336

RESULT 15
US-11-087-099-7483
; Sequence 7483, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 7483
; LENGTH: 983
; TYPE: PRT
; ORGANISM: Pseudomonas syringae pv. syringae B728a
US-11-087-099-7483

Query Match      12.9%; Score 60.5; DB 7; Length 983;
Best Local Similarity 33.3%; Pred. No. 33;
Matches 15; Conservative 9; Mismatches 10; Indels 11; Gaps 2;

Qy 2 SGGVKVQYKNNDSAPGD--NQIKPGLQ-----VNTGSSSV 35
Db 522 NAGVKAEIFSNTSFSGDPLTRVEFGVNLNWTGTVNAGTAV 566

Search completed: March 2, 2006, 14:28:14
Job time : 1.78442 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 2, 2006, 12:06:58 ; Search time 31.7439 Seconds
(without alignments)
4927.737 Million cell updates/sec

Title: US-09-917-376-5
Perfect score: 469
Sequence: 1 VSGGVKQVQKNDSPGDNQ.....TRAFSGVNPATPATDYIQ 88

Scoring table:
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/abs/ABSSWEB spool/US09917376/runat_02032006_091454_8101/app_query.fasta_1
-DB=Issued Patents_NA -QFMT=fastcap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blotsum62 -TRANS=human40.cdi
-MODE=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-HOST=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs06h -USER=US09917376 @CGN 1.1.278 @runat_02032006_091454_8101
-NCPU=6 -ICPU=3 -NO WMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents_NA:
1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	195.5	41.7	2029	3	US-09-136-574A-46
2	194.5	41.5	6416	3	US-09-136-574A-2
3	192.5	41.0	11707	2	US-09-136-574A-1
4	182.5	38.9	2977	2	US-07-862-588B-1
5	165	35.2	1438	3	US-09-339-159B-3
6	165	35.2	1482	3	US-09-198-956-9
7	165	35.2	1482	3	US-09-198-955A-11
8	165	35.2	1482	3	US-09-694-531-11
9	165	35.2	1482	3	US-09-670-141-9

10	165	35.2	1482	3	US-10-072-152-11	Sequence 11, Appl
11	155.5	33.2	4059	3	US-09-784-554B-1	Sequence 1, Appl
12	150.5	32.1	4056	3	US-09-784-554B-3	Sequence 3, Appl
13	120.5	25.7	1624	2	US-07-862-588B-6	Sequence 6, Appl
14	120.5	25.7	1775	2	US-07-862-588B-5	Sequence 5, Appl
15	114	24.3	486	2	US-08-048-164A-1	Sequence 1, Appl
16	114	24.3	486	2	US-08-048-164A-3	Sequence 3, Appl
17	114	24.3	486	2	US-08-460-462-1	Sequence 1, Appl
18	114	24.3	486	2	US-08-460-462-3	Sequence 3, Appl
19	114	24.3	486	2	US-08-460-457-1	Sequence 1, Appl
20	114	24.3	486	2	US-08-460-457-3	Sequence 3, Appl
21	114	24.3	486	2	US-08-460-458-1	Sequence 1, Appl
22	114	24.3	486	2	US-08-460-458-3	Sequence 3, Appl
23	114	24.3	486	2	US-08-460-455-1	Sequence 1, Appl
24	114	24.3	486	2	US-08-460-455-3	Sequence 3, Appl
25	114	24.3	486	2	US-08-330-394A-1	Sequence 1, Appl
26	114	24.3	486	2	US-08-330-394A-3	Sequence 3, Appl
27	114	24.3	499	3	US-09-006-636-5	Sequence 5, Appl
28	114	24.3	499	3	US-09-006-632-5	Sequence 5, Appl
29	114	24.3	499	3	US-09-325-274-5	Sequence 5, Appl
30	113	24.1	1146	3	US-09-277-716-21	Sequence 21, Appl
31	113	24.1	1146	3	US-09-609-161B-21	Sequence 21, Appl
32	78.5	16.7	764	3	US-09-533-559-5404	Sequence 5404, Ap
33	77	16.4	8107	3	US-09-335-586-3	Sequence 3, Appl
34	76	16.2	136058	3	US-09-949-016-12565	Sequence 12565, A
35	76	16.2	136480	3	US-09-949-016-17064	Sequence 17064, A
36	75.5	16.1	5714	3	US-09-620-312D-393	Sequence 393, App
37	72.5	15.5	819	3	US-09-902-540-3752	Sequence 3752, Ap
38	72.5	15.5	17897	3	US-09-902-540-1182	Sequence 1182, Ap
39	72.5	15.5	20284	3	US-09-526-193A-21	Sequence 21, Appl
40	70	14.9	10892	3	US-09-902-540-962	Sequence 962, App
41	70	14.9	25165	3	US-09-453-702B-39	Sequence 39, Appl
42	70	14.9	25165	3	US-10-114-170-39	Sequence 39, Appl
43	70	14.9	246444	3	US-09-949-016-13113	Sequence 13113, A
44	69.5	14.8	103987	3	US-09-949-016-12513	Sequence 12513, A
45	69.5	14.8	103988	3	US-09-949-016-17050	Sequence 17050, A

ALIGNMENTS

RESULT 1
US-09-136-574A-46
; Sequence 46, Application US/09136574A
; Patent No. 6294366

GENERAL INFORMATION:

APPLICANT: Farrington, Graham K.

Anderson, Paige

Gibbs, Moreland

Bergquist, Peter

Daniels, Roy

Morgan, Hugh W.

Williams, Diane P.

TITLE OF INVENTION: Compositions and Methods for Treating Cellulose Containing Fabrics Using Truncated Cellulase Enzyme Compositions

NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Center, P.O. Box 457

CITY: Spring House

STATE: PA

COUNTRY: USA

ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/136.574A

FILING DATE: 19-Aug-1998

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/932,571

ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Center, P.O. Box 457
CITY: Spring House
STATE: PA
COUNTRY: USA
ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/136,574A

FILING DATE: 19-Aug-1998

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/932,571

FILING DATE: September 19, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Bak, Mary E.

REGISTRATION NUMBER: 31,215

REFERENCE/DOCKET NUMBER: 1997US001/CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-540-9200

TELEFAX: 215-540-5818

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 11707 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-136-574A-1

Alignment Scores:
Pred. No.: 1,698-15 Length: 11707
Score: 192.50 Matches: 37
Percent Similarity: 62.4% Conservative: 16
Best Local Similarity: 43.5% Mismatches: 31
Query Match: 41.0% Indels: 1
DB: 3 Gaps: 1

US-09-917-376-5 (1-88) x US-09-136-574A-1 (1-11707)

QY 4 GlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnIleLysPro 23
DB 4038 GGTGAGGTACTATACAGACACATGAGCAAGTGCAGCACAAGTTCTATAAGCGC 4097
QY 24 GlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArg 43
DB 4098 TGGTTAAGATAGTAGTAATGGAGGACAGCAGTGTGATCTTAGCAGGTTAAGATAAGA 4157
QY 44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
DB 4158 TACTGGTACACAGTGGATGGTGTGACAAAGCCACAGAGTGCAGTA---TGTACTGGGCACAG 4214
QY 64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
DB 4215 ATAGGGGCAACCAATGTGACATTCATTTTGTGAAGCTGAGCAGCGGAGTGAGTGAGCG 4274
QY 84 AspThrTyrLeuGln 88
DB 4275 GATTATTACTTGGAG 4289

RESULT 4

US-07-862-588B-1

Sequence 1, Application US/07862588B

Patent No. 5916796

GENERAL INFORMATION:

APPLICANT: Joergensen, Per Linea

APPLICANT: Sch lein, Martin

APPLICANT: Hansen, Christian

TITLE OF INVENTION: An Enzyme Exhibiting Cellulase Activity

NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59167960 No. 5916796disk of No. 5916796th America, Inc.
STREET: 405 Lexington Avenue, 62nd floor
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10017
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/862,588B
FILING DATE: 19920727
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DK 164/90
FILING DATE: 19-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/DK91/00013
FILING DATE: 18-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Zelson, Steve T. / Lambiris, Elias J.
REGISTRATION NUMBER: 30,335 / 33,728
REFERENCE/DOCKET NUMBER: 3425.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2977 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Bacillus lautus
STRAIN: NCIMB 40250
FEATURE:
NAME/KEY: CDS
LOCATION: 677..2776
OTHER INFORMATION:
US-07-862-588B-1

Alignment Scores:
Pred. No.: 5,76e-15 Length: 2977
Score: 182.50 Matches: 36
Percent Similarity: 66.3% Conservative: 19
Best Local Similarity: 43.4% Mismatches: 25
Query Match: 38.9% Indels: 3
DB: 2 Gaps: 2

US-09-917-376-5 (1-88) x US-07-862-588B-1 (1-2977)

QY 6 LysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnIleLysProGlyLeu 25
DB 144 CGACTGCATACAGCGCGCCGATACAAATGACGAGCAGCAACAGATCAAGCGCTCCTTC 203
QY 26 GlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArgTyrTrp 45
DB 204 AACATCAAAACACACGCTACTTCGGCTGTGATTATTAAGCAGCGTCAAAATCCGCTACTAC 263
QY 46 PheThrArgAspGlyGlySerThrLeuValTyrAsnCysAspTrpAlaAlaMetGly 65
DB 264 TTCACCAAGATGGTTCTCGCGCGGTGAACGGCTGG---ATCGACTGGGCGCAGCTCGGC 320
QY 66 CysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAlaAspThr 85
DB 321 GGCAGCAACATTCAGATCTCGTTTGGC-----AACCACTACTGGCAGCAATTCGATACG 374
QY 86 TyrLeuGln 88

```
|||||: 375 TACGTGGG 383
; EARLIER FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 1482
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-198-956-9

Alignment Scores:
Pred. No.: 5,788-13 Length: 1482
Score: 165.00 Matches: 35
Percent Similarity: 59.8% Conservative: 23
Best Local Similarity: 36.1% Mismatches: 29
Query Match: 35.2% Indels: 10
DB: 3 Gaps: 2

US-09-917-376-5 (1-88) x US-09-198-956-9 (1-1482)
QY 1 ValSerGlyValLeuValGlnThrValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 991 GTATCAGGCAATTGGAAGTTGAATCTACAACAGCAATCTTCAGATACTACTACTCA 1050
QY 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 1051 ATCAATCCTCAGTTCAGGTTACTAATACCGAAGCAGTCAATTGATTGTCCAAATC 1110
QY 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
Db 1111 ACATTGAGATATTATTATACAGTAGACGACAGAAAGATCAGACCTTCTGG---TGTGAC 1167
QY 61 TrpAlaAlaMet-----GlyCysGlyAsnIleArgAla 71
Db 1168 CATGCTGCAATAATCGGCAGTAACGCGAGCTACAACGGAATTACTTCAATGTAAAGGA 1227
QY 72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
Db 1228 ACATTGTAAAAATGAGTTCTCTCAACAAATAACGACAGACACCTACCTTGAA 1278

RESULT 7
US-09-198-955A-11
; Sequence 11, Application US/09198955A
; Patent No. 6187580
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schulein, Martin
; APPLICANT: Lange, Niels E.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Moller, Soren
; APPLICANT: Glad, Sanne O. S.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Schnorr, Kirk
; APPLICANT: Kongsbak, Lars
; TITLE OF INVENTION: No. 6187580el Pectate Lyases
; FILE REFERENCE: 5378.200-US
; CURRENT APPLICATION NUMBER: US/09/198,955A
; PRIOR FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: 1343/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 1344/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/067,249
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 60/067,240
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 09/073,684
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 09/184,217
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1482

Db 375 TACGTGGG 383
RESULT 5
US-09-339-159B-3
; Sequence 3, Application US/09339159B
; Patent No. 6566114
; GENERAL INFORMATION:
; APPLICANT: Kauppinen, Markus
; APPLICANT: Schulein, Martin
; APPLICANT: Schnorr, Kirk
; APPLICANT: Andersen, Lene
; APPLICANT: Bjornvad, Mads
; TITLE OF INVENTION: No. 6566114el Mannanases
; FILE REFERENCE: 5440.204-US
; CURRENT APPLICATION NUMBER: US/09/339,159B
; CURRENT FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 1438
; TYPE: DNA
; ORGANISM: Bacillus sp. 1633
US-09-339-159B-3

Alignment Scores:
Pred. No.: 5,538-13 Length: 1438
Score: 165.00 Matches: 35
Percent Similarity: 59.8% Conservative: 23
Best Local Similarity: 36.1% Mismatches: 29
Query Match: 35.2% Indels: 10
DB: 3 Gaps: 2

US-09-917-376-5 (1-88) x US-09-339-159B-3 (1-1438)
QY 1 ValSerGlyValLeuValGlnThrValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 940 GTATCAGGCAATTGGAAGTTGAATCTACAACAGCAATCTTCAGATACTACTACTCA 999
QY 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 1000 ATCAATCCTCAGTTCAGGTTACTAATACCGAAGCAGTCAATTGATTGTCCAAATC 1059
QY 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
Db 1060 ACATTGAGATATTATTATACAGTAGACGACGACAGAAAGATCAGACCTTCTGG---TGTGAC 1116
QY 61 TrpAlaAlaMet-----GlyCysGlyAsnIleArgAla 71
Db 1117 CATGCTGCAATAATCGGCAGTAACGCGAGCTACAACGGAATTACTTCAATGTAAAGGA 1176
QY 72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
Db 1177 ACATTGTAAAAATGAGTTCTCTCAACAAATAACGACAGACACCTACCTTGAA 1227

RESULT 6
US-09-198-956-9
; Sequence 9, Application US/09198956
; Patent No. 6165769
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schulein, Martin
; APPLICANT: Lange, Niels Erik K.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Schnorr, Kirk
; TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus
; FILE REFERENCE: 5377.200-US
; CURRENT APPLICATION NUMBER: US/09/198,956
; CURRENT FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: 1344/97
; EARLIER FILING DATE: 1997-11-24
; EARLIER APPLICATION NUMBER: 60/067,240
```

TYPE: DNA
ORGANISM: Clostridium thermocellum
US-09-198-955A-11

Alignment Scores:
Pred. No.: 5,78e-13 Length: 1482
Score: 165.00 Matches: 35
Percent Similarity: 59.8% Conservative: 23
Best Local Similarity: 36.1% Mismatches: 29
Query Match: 35.2% Indels: 10
DB: 3 Gaps: 2

US-09-917-376-5 (1-88) x US-09-198-955A-11 (1-1482)

QY 1 ValSerGlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGln 20
Db 991 GTATCAGCAATTGAAGTTGAATTTCTACACAGCAATCCTTCACTACTACTCA 1050
QY 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 1051 ATCAATCCTCAGTTCAGGTTACTAATACCGAAGCAGTGCATTTGTTCCAACTC 1110
QY 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
Db 1111 ACATTGAGATATTATTATACAGTAGACGACGAGAAGATCAGACCTTCTGG---TGTGAC 1167
QY 61 TrpAlaAlaMet-----GlyCysGlyAsnIleArgAla 71
Db 1168 CATGCTGCAATAATCGGCAGTACCGCAGCTACACGGAATTACTTCAAAATGTAAGA 1227
QY 72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
Db 1228 ACATTGTAAAATGAGTTCTCTCAACAATAACGACGACACCTTCTTGAA 1278

RESULT 8

US-09-694-531-11
Sequence 11, Application US/09694531
Patent No. 6368843
GENERAL INFORMATION:
APPLICANT: Andersen, Lene N.
APPLICANT: Schulein, Martin
APPLICANT: Lange, Niels E.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Moller, Soren
APPLICANT: Glad, Sanne O. S.
APPLICANT: Kauppinen, Markus S.
APPLICANT: Schnorr, Kirk
APPLICANT: Kongsbak, Lars
TITLE OF INVENTION: No. 6368843el Pectate Lyases
FILE REFERENCE: 5378.200-US
CURRENT APPLICATION NUMBER: US/09/694,531
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 09/198,955
PRIOR FILING DATE: 1998-11-24
PRIOR APPLICATION NUMBER: 1343/97
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 1344/97
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/067,249
PRIOR FILING DATE: 1997-12-02
PRIOR APPLICATION NUMBER: 60/067,240
PRIOR FILING DATE: 1997-12-02
PRIOR APPLICATION NUMBER: 09/073,684
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 09/184,217
PRIOR FILING DATE: 1998-11-02
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 1482

TYPE: DNA
ORGANISM: Clostridium thermocellum
US-09-694-531-11

Alignment Scores:

Pred. No.: 5,78e-13 Length: 1482
Score: 165.00 Matches: 35
Percent Similarity: 59.8% Conservative: 23
Best Local Similarity: 36.1% Mismatches: 29
Query Match: 35.2% Indels: 10
DB: 3 Gaps: 2

US-09-917-376-5 (1-88) x US-09-694-531-11 (1-1482)

QY 1 ValSerGlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGln 20
Db 991 GTATCAGCAATTGAAGTTGAATTTCTACACAGCAATCCTTCACTACTACTCA 1050
QY 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 1051 ATCAATCCTCAGTTCAGGTTACTAATACCGAAGCAGTGCATTTGTTCCAACTC 1110
QY 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
Db 1111 ACATTGAGATATTATTATACAGTAGACGACGAGAAGATCAGACCTTCTGG---TGTGAC 1167
QY 61 TrpAlaAlaMet-----GlyCysGlyAsnIleArgAla 71
Db 1168 CATGCTGCAATAATCGGCAGTACCGCAGCTACACGGAATTACTTCAAAATGTAAGA 1227
QY 72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
Db 1228 ACATTGTAAAATGAGTTCTCTCAACAATAACGACGACACCTTCTTGAA 1278

RESULT 9

US-09-670-141-9
Sequence 9, Application US/09670141
Patent No. 6429000
GENERAL INFORMATION:
APPLICANT: Andersen, Lene N.
APPLICANT: Schulein, Martin
APPLICANT: Lange, Niels Erik K.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Schnorr, Kirk
TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus
TITLE OF INVENTION: Licheniformis
FILE REFERENCE: 5377.200-US
CURRENT APPLICATION NUMBER: US/09/670,141
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: 09/198,956
PRIOR FILING DATE: 1998-11-24
PRIOR APPLICATION NUMBER: 1344/97
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/067,240
PRIOR FILING DATE: 1997-12-02
NUMBER OF SEQ ID NOS: 26
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 9
LENGTH: 1482
TYPE: DNA
ORGANISM: Bacillus licheniformis
US-09-670-141-9

Alignment Scores:

Pred. No.: 5,78e-13 Length: 1482
Score: 165.00 Matches: 35
Percent Similarity: 59.8% Conservative: 23
Best Local Similarity: 36.1% Mismatches: 29
Query Match: 35.2% Indels: 10
DB: 3 Gaps: 2

US-09-917-376-5 (1-88) x US-09-670-141-9 (1-1482)

QY 1 ValSerGlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGln 20
Db 991 GTATCAGCAATTGAAGTTGAATTTCTACACAGCAATCCTTCACTACTACTCA 1050

```
Qy 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 1051 ATCAATCCTCAGTTCAAGTTACTAATACCGAAGCAGTGCATTAATGTTGTCAAACCTC 1110
Qy 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerThrLeuValTyrAsnCysAsp 60
Db 1111 ACATTGAGATATTATTATACAGTACGACGACAGAAAGATCAGACCTTCTCG---TGTGAC 1167
Qy 61 TrpAlaAlaMet-----GlyCysGlyAsnIleArgAla 71
Db 1168 CATGCTGCAATAATCGGCAGTAACGGCAGCTACACGGAATTACTTCAAAATGTAAGA 1227
Qy 72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
Db 1228 ACATTGTGAAAATGAGTTCCTCAACAAATAACGACAGACCTACCTTGAA 1278
RESULT 10
US-10-072-152-11
; Sequence 11, Application US/10072152
; Patent No. 6677147
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schulein, Martin
; APPLICANT: Lange, Niels E.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Moller, Soren
; APPLICANT: Glad, Sanne O. S.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Schnorr, Kirk
; APPLICANT: Kongsbak, Lars
; TITLE OF INVENTION: No. 6677147el Pectate Lyases
; FILE REFERENCE: 5378.200-US
; CURRENT APPLICATION NUMBER: US/10/072,152
; CURRENT FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: US/09/198,955
; PRIOR FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: 1343/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 1344/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/067,249
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 60/067,240
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 09/073,684
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 09/184,217
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1482
; TYPE: DNA
; ORGANISM: Clostridium thermocellum
US-10-072-152-11
Alignment Scores:
Pred. No.: 5.78e-13 Length: 1482
Score: 165.00 Matches: 35
Percent Similarity: 59.8% Conservative: 23
Best Local Similarity: 36.1% Mismatches: 29
Query Match: 35.2% Indels: 10
DB: 3 Gaps: 2
US-09-917-376-5 (1-88) x US-10-072-152-11 (1-1482)
Qy 1 ValSerGlyValValGlnTyrLysAsnAspSerAlaProGlyAspAsnGln 20
Db 991 GTATCGCAATTTGAGAGTTGAATTTACAAAGCAAGTCTTCAGATACCTACTACTCA 1050
Qy 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 1111 ACATTGAGATATTATTATACAGTACGACGACAGAAAGATCAGACCTTCTCG---TGTGAC 1167
```

```
Db 1051 ATCAATCCTCAGTTCAAGTTACTAATACCGAAGCAGTGCATTAATGTTGTCAAACCTC 1110
Qy 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerThrLeuValTyrAsnCysAsp 60
Db 1111 ACATTGAGATATTATTATACAGTACGACGACAGAAAGATCAGACCTTCTCG---TGTGAC 1167
Qy 61 TrpAlaAlaMet-----GlyCysGlyAsnIleArgAla 71
Db 1168 CATGCTGCAATAATCGGCAGTAACGGCAGCTACACGGAATTACTTCAAAATGTAAGA 1227
Qy 72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
Db 1228 ACATTGTGAAAATGAGTTCCTCAACAAATAACGACAGACCTACCTTGAA 1278
RESULT 11
US-09-784-554B-1
; Sequence 1, Application US/09784554B
; Patent No. 6815192
; GENERAL INFORMATION:
; APPLICANT: Schnorr, Kirk
; APPLICANT: Jorgensen, Per Lina
; APPLICANT: Schulein, Martin
; TITLE OF INVENTION: FAMILY 44 XYLOGLUCANASES
; FILE REFERENCE: 10017.200-US
; CURRENT APPLICATION NUMBER: US/09/784,554B
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4059
; TYPE: DNA
; ORGANISM: Paenibacillus polymyxa
US-09-784-554B-1
Alignment Scores:
Pred. No.: 5.25e-11 Length: 4059
Score: 155.50 Matches: 27
Percent Similarity: 56.5% Conservative: 21
Best Local Similarity: 31.8% Mismatches: 36
Query Match: 33.2% Indels: 1
DB: 3 Gaps: 1
US-09-917-376-5 (1-88) x US-09-784-554B-1 (1-4059)
Qy 4 GlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnIleLysPro 23
Db 3598 GGACTGCTGCTCCAATATCGCACAGATATACTAGGTGAACGATAATCACCTCAATCCG 3657
Qy 24 GlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValValArg 43
Db 3658 CAATTCCAAATTTGTAACAAGGCACACCTCCATACCGATCAACAGTTGAAATTCGC 3717
Qy 44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
Db 3718 TACTACTACACAATCGACGGTGCACCGTGCAGCAG---ACTTTCACCTGCAGCATTCGCACG 3774
Qy 64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
Db 3775 CTGAGCTGCTCAAAGCTGAACGGTAACTGGTTAAATGAGAGGCTGCACCGGTGCC 3834
Qy 84 AspThrTyrLeuGln 88
Db 3835 GATTATTATTGGA 3849
RESULT 12
US-09-784-554B-3
; Sequence 3, Application US/09784554B
; Patent No. 6815192
; GENERAL INFORMATION:
; APPLICANT: Schnorr, Kirk
; APPLICANT: Jorgensen, Per Lina
; APPLICANT: Schulein, Martin
; TITLE OF INVENTION: FAMILY 44 XYLOGLUCANASES
```

```
; FILE REFERENCE: 10017.200-US
; CURRENT APPLICATION NUMBER: US/09/784, 554B
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 4056
; TYPE: DNA
; ORGANISM: Paenibacillus polymyxa
US-09-784-554B-3

Alignment Scores:
Pred. No.: 2,61e-10 Length: 4056
Score: 150.50 Matches: 27
Percent Similarity: 56.5% Conservative: 21
Best Local Similarity: 31.8% Mismatches: 36
Query Match: 32.1% Indels: 1
DB: 3 Gaps: 1

US-09-917-376-5 (1-88) x US-09-784-554B-3 (1-4056)
QY 4 GlyValValValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnIleLysPro 23
Db 3595 GGATTGGTCTCCAGTATCGCACGGGATCAAAATGTGACGACAACTCACTGAACCG 3654
QY 24 GlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArg 43
Db 3655 CATTTCCAAATTTAAATAAAGGTACATCTCCGTACCGATCAACGAGTTGAAATTCG 3714
QY 44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnGlyAspTrpAla 63
Db 3715 TACTACTACGATCGACGGTGCACCGTGAGCAG---ACATTCACTGCGACTATCGGTG 3771
QY 64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
Db 3772 CTGACTGCTCGAGCTGAATGTAAGTGTAAATGATAAAGCTGCAACCGGTGCT 3831
QY 84 AspThrTyrLeuGln 88
Db 3832 GATTATTATTGGAA 3846

RESULT 13
US-07-862-588B-6
; Sequence 6, Application US/07862588B
; Patent No. 5916796
; GENERAL INFORMATION:
; APPLICANT: Joergensen, Per Linnae
; APPLICANT: Sch lein, Martin
; APPLICANT: Hansen, Christian
; TITLE OF INVENTION: An Enzyme Exhibiting Cellulase Activity
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 59167960 No. 5916796disk of No. 5916796th America, Inc.
; STREET: 405 Lexington Avenue, 62nd floor
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/862,588B
; FILING DATE: 19920727
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 164/90
; FILING DATE: 19-JAN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/DK91/00013
; FILING DATE: 18-JAN-1991

; ATTORNEY/AGENT INFORMATION:
; NAME: Zelson, Steve T. / Lambiris, Elias J.
; REGISTRATION NUMBER: 30,335 / 33,728
; REFERENCE/DOCKET NUMBER: 3425.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1624 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Bacillus lautus
; STRAIN: NCIMB 40250
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 30..1607
; OTHER INFORMATION:
US-07-862-588B-6

Alignment Scores:
Pred. No.: 1.06e-06 Length: 1624
Score: 120.50 Matches: 28
Percent Similarity: 54.1% Conservative: 18
Best Local Similarity: 32.9% Mismatches: 36
Query Match: 25.7% Indels: 3
DB: 2 Gaps: 2

US-09-917-376-5 (1-88) x US-07-862-588B-6 (1-1624)
QY 2 SerGlyValValValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnIle 21
Db 1323 ACGGGAACTTGTTCCTCAATAACAAGTTGGGACACTAGCGCCACGATAACCAATG 1382
QY 22 LysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThr 41
Db 1383 AAGCTTCTCTTAACATCAAGAACACCGTACAAACCCCTGTTAACTTGAGCGGCTCAAG 1442
QY 42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnGlyAspTrp 61
Db 1443 CTTNNNNNNNNNNNNAAAAAGAC---GGACCTGGCGGATATGAGTCTCGATCGACTGG 1499
QY 62 AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro 81
Db 1500 GCGCAATCGCGCAACGAATGTTCTGCTGGCATTG-----GCTAACTTTACCGGGAGT 1553
QY 82 ThrAlaAspThrTyr 86
Db 1554 AATACGGATACTTAC 1568

RESULT 14
US-07-862-588B-5
; Sequence 5, Application US/07862588B
; Patent No. 5916796
; GENERAL INFORMATION:
; APPLICANT: Joergensen, Per Linnae
; APPLICANT: Sch lein, Martin
; APPLICANT: Hansen, Christian
; TITLE OF INVENTION: An Enzyme Exhibiting Cellulase Activity
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 59167960 No. 5916796disk of No. 5916796th America, Inc.
; STREET: 405 Lexington Avenue, 62nd floor
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```



```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/862,588B
; FILING DATE: 19920727
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 164/90
; FILING DATE: 19-JAN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/DK91/00013
; FILING DATE: 18-JAN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Zelson, Steve T. / Lambiris, Elias J.
; REGISTRATION NUMBER: 30,335 / 33,728
; REFERENCE/DOCKET NUMBER: 3425.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1775 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Bacillus lautus
; STRAIN: NCIMB 40250
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 30..(1625,1775)
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /partial
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /trans_except= (pos: 1446 .. 1458, aa:
; OTHER INFORMATION: OTR)
US-07-862-588B-5

```

```

Alignment Scores:
Pred. No.: 1,28-06 Length: 1775
Score: 120.50 Matches: 28
Percent Similarity: 54.1% Conservative: 18
Best Local Similarity: 32.9% Mismatches: 36
Query Match: 25.7% Indels: 3
DB: 2 Gaps: 2

```

US-09-917-376-5 (1-88) x US-07-862-588B-5 (1-1775)

```

Qy 2 SerGlyGlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnIle 21
Db 1323 ACGGGGAACCTGTGTGTCACATCAACAGCTGCGGACACTGCGGACCGGATAACAAATG 1382

Qy 22 LysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThr 41
Db 1383 AAGCTTCCTTTACATCAAGACACGATCAACCCCTGTTACCTGAGCGGCTCAAG 1442

Qy 42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp 61
Db 1443 CTTNNNNNNNNNNNNAAAAAGAC--GGACCTGGGATATGAGTGTCTGATCGCTG 1499

Qy 62 AlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro 81
Db 1500 GCGCAATCGGCGCAAGCAATGTTCTGTGTCGCAATTC-----GCTAACTTTACCGGGAGT 1553

Qy 82 ThrAlaAspThrTyr 86
Db 1554 AATACGGATACCTTAC 1568

RESULT 15
US-08-048-164A-1

```

```

; Sequence 1, Application US/08048164A
; Patent No. 5496934
; GENERAL INFORMATION:
; APPLICANT: Shoseyov, Oded
; APPLICANT: Shiegl, Itai
; APPLICANT: Goldstein, Marc A.
; APPLICANT: Doi, Roy H.
; TITLE OF INVENTION: CELLULOSE BINDING DOMAIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/048,164A
; FILING DATE: 14-APR-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7809-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 486 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..486
US-08-048-164A-1

Alignment Scores:
Pred. No.: 1,49e-06 Length: 486
Score: 114.00 Matches: 29
Percent Similarity: 52.6% Conservative: 22
Best Local Similarity: 29.9% Mismatches: 34
Query Match: 24.3% Indels: 12
DB: 2 Gaps: 4

US-09-917-376-5 (1-88) x US-08-048-164A-1 (1-486)

Qy 2 SerGlyGlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnIle 21
Db 7 ACATCATCAATGTCAGTTGAATTTTACAACTCTAACAAATTCAGCAACAACTCAATT 66

Qy 22 LysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThr 41
Db 67 ACACCAATAATCAAAATTAATTAACATCTGACAGTGATTAAATTTAAATGACGTAAAA 126

Qy 42 ValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrp 61
Db 127 GTTAGATATTATTACCAAGTGTGATGTCACACAGGACAACTTCTCGG---TGTGACCAT 183

Qy 62 AlaAlaMetGlyCysGlyAsn-----IleArgAlaSerPhe 73
Db 184 GCTGTGTCATTATTAGGAATAGCTATGTTGATAACACTAGCAAAAGTACGAACTTC 243

Qy 74 -----GlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88

```

Db 244 GTTAAAGAAACAGCAAGCCCA---ACATCAACCTATGATACATATGTTGAA 291

Search completed: March 2, 2006, 13:50:35
Job time : 36.7439 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 2, 2006, 12:19:26 ; Search time 165.153 Seconds
(without alignments)
4406.259 Million cell updates/sec

Title: US-09-917-376-5

Perfect score: 469

Sequence: 1 VSGGVKQVKNDSAPGDNQ.....IRASFGVNPATPTADTYLQ 88

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEVEL=slp
-Q=/abs/ABSWEBSPOOL/US09917376/runat_02032006_091500_8223/app_query.fasta_1
-DB=Published Applications NA Main -QFMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs06p
-USER=US09917376@CGN_1_1_3095@runat_02032006_091500_8223 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	469	100.0	2289	3	US-09-917-378-2
2	466	99.4	3687	3	US-09-917-384-2
3	466	99.4	3687	3	US-09-917-383-2
4	462	98.5	2869	3	US-09-917-376-2
5	462	98.5	2869	3	US-10-155-400-2
6	257	54.8	2223	6	US-10-156-761-550
7	257	54.8	9025608	6	US-10-156-761-1

8	190	40.5	1527	6	US-10-369-493-46838	Sequence 46838, A
9	188	40.1	1434	8	US-10-466-208-11	Sequence 11, Appl
10	188	40.1	1491	8	US-10-466-208-7	Sequence 7, Appl
11	188	40.1	2510	7	US-10-433-577-34	Sequence 34, Appl
12	165	35.2	501	9	US-10-933-404-1	Sequence 1, Appl
13	165	35.2	1438	6	US-10-372-054-3	Sequence 3, Appl
14	165	35.2	1482	5	US-10-072-152-11	Sequence 11, Appl
15	165	35.2	1482	7	US-10-855-433-11	Sequence 11, Appl
16	157	33.5	1314	3	US-09-974-300-684	Sequence 684, App
17	155.5	33.2	4059	3	US-09-784-554B-1	Sequence 1, Appl
18	155.5	33.2	4059	3	US-10-896-555-1	Sequence 1, Appl
19	150.5	32.1	4056	3	US-09-784-554B-3	Sequence 3, Appl
20	150.5	32.1	4056	8	US-10-896-555-3	Sequence 3, Appl
21	131.5	28.0	3180	9	US-10-886-393A-29	Sequence 29, Appl
22	130.5	27.8	4452	7	US-10-282-122A-15299	Sequence 15299, A
23	126.5	27.0	1303	9	US-10-490-737-1	Sequence 1, Appl
24	126.5	27.0	1747	9	US-10-490-737-2	Sequence 2, Appl
25	126.5	27.0	3489	9	US-10-490-737-3	Sequence 3, Appl
26	114	24.3	768	6	US-10-261-446-5	Sequence 5, Appl
27	114	24.3	768	7	US-10-261-446B-5	Sequence 9, Appl
28	113	24.1	652	8	US-10-782-234-5	Sequence 5, Appl
29	113	24.1	652	8	US-10-748-055-9	Sequence 9, Appl
30	113	24.1	1146	3	US-09-808-898-21	Sequence 21, Appl
31	78.5	16.7	764	8	US-10-853-047-5404	Sequence 5404, Ap
32	78.5	16.7	3147	7	US-10-282-122A-13716	Sequence 13716, A
33	76	16.2	757	7	US-10-767-701-8547	Sequence 8547, Ap
34	76	16.2	96597	7	US-10-052-482-226	Sequence 236, App
35	75.5	16.1	1041	7	US-10-260-238-742	Sequence 742, App
36	75.5	16.1	1041	7	US-10-437-963-76751	Sequence 76751, A
37	75.5	16.1	2091	7	US-10-211-462-222	Sequence 222, App
38	75.5	16.1	2091	8	US-10-723-860-4038	Sequence 4038, Ap
39	75.5	16.1	2091	9	US-10-756-149-4006	Sequence 4006, Ap
40	75.5	16.1	3396	7	US-10-398-458-1	Sequence 1, Appl
41	75.5	16.1	4041	3	US-09-747-835A-21	Sequence 21, Appl
42	75.5	16.1	4041	7	US-10-312-312-21	Sequence 21, Appl
43	75.5	16.1	4353	5	US-10-225-567A-405	Sequence 405, App
44	75.5	16.1	4755	9	US-10-505-486-201	Sequence 201, App
45	75.5	16.1	4914	3	US-09-974-298-188	Sequence 188, App

ALIGNMENTS

RESULT 1

US-09-917-378-2
; Sequence 2, Application US/09917378
; Publication NO. US20030119093A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT MANNANASE FROM ACIDOTHERMUS CELLULOLYTICUS
; FILE REFERENCE: 40197.7US01
; CURRENT APPLICATION NUMBER: US/09/917,378
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2289
; TYPE: DNA
; ORGANISM: Acidothermus cellulolyticus
US-09-917-378-2

Alignment Scores:	
Pred. No.:	6,948-56
Score:	469.00
Percent Similarity:	100.0%
Best Local Similarity:	100.0%
Query Match:	100.0%
DB:	3
Length:	2289
Matches:	88
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0


```
; CURRENT APPLICATION NUMBER: US/09/917,376
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2869
; TYPE: DNA
; ORGANISM: Acidothermus cellulosoliticus
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (2869)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-09-917-376-2

Alignment Scores:
Pred. No.: 9,34e-55 Length: 2869
Score: 462.00 Matches: 86
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 97.7% Mismatches: 0
Query Match: 98.5% Indels: 0
DB: 3 Gaps: 0

US-09-917-376-5 (1-88) x US-09-917-376-2 (1-2869)

Qy 1 ValSerGlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGln 20
Db 2605 GTGTCGGGTGGGTGAAGTGCAGTATAGAAATAATGATTCGGCGCCGGTGATAATCAG 2664

Qy 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 2665 ATCAAGCCGGTTCAGAGTGGTGAATACCGGCTCGTTCGGCTCGGTGAATTTGTCACGGTG 2724

Qy 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
Db 2725 ACGTTCGGGTACTGGTTCACCCGGGATGGTGGCTCGACACTGGTGTACACTGTGAC 2784

Qy 61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
Db 2785 TGGGGCGGATCGGTGTGGGAATATCCGCGCTCGTTCGGCTCGGTGAACCCGCGCAGC 2844

Qy 81 ProThrAlaAspThrTyrLeuGln 88
Db 2845 CCGACGGCGGACACTACCTGCAG 2868

RESULT 5
US-10-155-400-2
; Sequence 2, Application US/10155400
; Publication No. US2003010988A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT AVICELASE FROM ACIDOTHERMUS
; FILE REFERENCE: NREL 01-36A
; CURRENT APPLICATION NUMBER: US/10/155,400
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2869
; TYPE: DNA
; ORGANISM: Acidothermus cellulosoliticus
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (2869)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-10-155-400-2

Alignment Scores:
Pred. No.: 9,34e-55 Length: 2869
Score: 462.00 Matches: 86

US-09-917-376-5 (1-88) x US-10-155-400-2 (1-2869)

Qy 1 ValSerGlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGln 20
Db 2605 GTGTCGGGTGGGTGAAGTGCAGTATAGAAATAATGATTCGGCGCCGGTGATAATCAG 2664

Qy 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 2665 ATCAAGCCGGTTCAGAGTGGTGAATACCGGCTCGTTCGGCTCGGTGAATTTGTCACGGTG 2724

Qy 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
Db 2725 ACGTTCGGGTACTGGTTCACCCGGGATGGTGGCTCGACACTGGTGTACACTGTGAC 2784

Qy 61 TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThr 80
Db 2785 TGGGGCGGATCGGTGTGGGAATATCCGCGCTCGTTCGGCTCGGTGAACCCGCGCAGC 2844

Qy 81 ProThrAlaAspThrTyrLeuGln 88
Db 2845 CCGACGGCGGACACTACCTGCAG 2868

RESULT 6
US-10-156-761-550
; Sequence 550, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 550
; LENGTH: 2223
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) .. (2223)
US-10-156-761-550

Alignment Scores:
Pred. No.: 1.17e-25 Length: 2223
Score: 257.00 Matches: 49
Percent Similarity: 71.3% Conservative: 13
Best Local Similarity: 56.3% Mismatches: 25
Query Match: 54.8% Indels: 0
DB: 6 Gaps: 0

US-09-917-376-5 (1-88) x US-10-156-761-550 (1-2223)

Qy 2 SerGlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGln 21
Db 1771 TCGGGCGGCTCAAGGTCTCTACAGAACACAGACTCTCTCGGCCACCGACCGCATC 1830

Qy 22 LysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThr 41
Db 22 LysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThr 41
```



```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CMCase gene with additional his encoding sequences
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1431)
US-10-466-208-11
Alignment Scores:
Pred. No.: 4,53e-16 Length: 1434
Score: 188.00 Matches: 34
Percent Similarity: 61.2% Conservative: 18
Best Local Similarity: 40.0% Mismatches: 31
Query Match: 40.1% Indels: 2
DB: 8 Gaps: 1

US-09-917-376-5 (1-88) x US-10-466-208-11 (1-1434)
QY 4 GlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnIleLysPro 23
Db 991 GGTATTCTGTACATAACAGCAGCGGATGGAGTATGAACAGCAACCAATCCGTCG 1050
QY 24 GlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArg 43
Db 1051 CAGCTTCAATATAAATAACCGCAATACCGGTTGATTAAAGATGTCACTGCCCGT 1110
QY 44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
Db 1111 TACTGGTATACGCGAAACAAAGGCCAAACGTT-----GACTGTGACTACGCGCAG 1164
QY 64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
Db 1165 CTGGATCGCGCAATGTGACATACAAAGTTTGTGACGTTGCTAAACCAACCAAGAGTGCA 1224
QY 84 AspThrTyrLeuGln 88
Db 1225 GATACCTATCTGGAA 1239

RESULT 10
US-10-466-208-7
; Sequence 7, Application US/10466208
; Publication No. US20040180348A1
; GENERAL INFORMATION:
; APPLICANT: GENOFOCUS Co., Ltd.
; TITLE OF INVENTION: Method for Surface Display of Proteins on Genetic Carriers
; FILE REFERENCE: 3260-13
; CURRENT APPLICATION NUMBER: US/10/466,208
; CURRENT FILING DATE: 2003-07-14
; PRIOR APPLICATION NUMBER: KR2001-2156
; PRIOR FILING DATE: 2001-01-15
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: KopatentIn 1.71
; SEQ ID NO 7
; LENGTH: 1491
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CMCase gene with mutated signal sequence to enhance its
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1488)
US-10-466-208-7
Alignment Scores:
Pred. No.: 4,77e-16 Length: 1491
Score: 188.00 Matches: 34
Percent Similarity: 61.2% Conservative: 18
Best Local Similarity: 40.0% Mismatches: 31
Query Match: 40.1% Indels: 2
DB: 8 Gaps: 1

US-09-917-376-5 (1-88) x US-10-466-208-7 (1-1491)
QY 4 GlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnIleLysPro 23
Db 2051 GGTATTCTGTACATAACAGCAGCGGATGGAGTATGAACAGCAACCAATCCGTCG 2110
QY 24 GlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArg 43
Db 2111 CAGCTTCAATATAAATAACCGCAATACCGGTTGATTAAAGATGTCACTGCCCGT 2170
QY 44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
Db 2171 TACTGGTATACGCGAAACAAAGGCCAAACGTT-----GACTGTGACTACGCGCAG 2224
QY 64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
Db 2225 CTGGATCGCGCAATGTGACATACAAAGTTTGTGACGTTGCTAAACCAACCAAGAGTGCA 2284
QY 84 AspThrTyrLeuGln 88
Db 2285 GATACCTATCTGGAA 2299

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CMCase gene with additional his encoding sequences
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1431)
US-10-466-208-11
Alignment Scores:
Pred. No.: 4,53e-16 Length: 1434
Score: 188.00 Matches: 34
Percent Similarity: 61.2% Conservative: 18
Best Local Similarity: 40.0% Mismatches: 31
Query Match: 40.1% Indels: 2
DB: 8 Gaps: 1

US-09-917-376-5 (1-88) x US-10-466-208-11 (1-1434)
QY 4 GlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnIleLysPro 23
Db 1048 GGTATTCTGTACATAACAGCAGCGGATGGAGTATGAACAGCAACCAATCCGTCG 1107
QY 24 GlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArg 43
Db 1108 CAGCTTCAATATAAATAACCGCAATACCGGTTGATTAAAGATGTCACTGCCCGT 1167
QY 44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
Db 1168 TACTGGTATACGCGAAACAAAGGCCAAACGTT-----GACTGTGACTACGCGCAG 1221
QY 64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
Db 1222 CTGGATCGCGCAATGTGACATACAAAGTTTGTGACGTTGCTAAACCAACCAAGAGTGCA 1281
QY 84 AspThrTyrLeuGln 88
Db 1282 GATACCTATCTGGAA 1296

RESULT 11
US-10-433-577-34
; Sequence 34, Application US/10433577
; Publication No. US20040171065A1
; GENERAL INFORMATION:
; APPLICANT: GENOFOCUS Co., Ltd.
; TITLE OF INVENTION: Method for Expression of Proteins on Spore Surface
; FILE REFERENCE: PCT-Genofocus-1
; CURRENT APPLICATION NUMBER: US/10/433,577
; CURRENT FILING DATE: 2003-06-05
; PRIOR APPLICATION NUMBER: KR2000-74835
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: KopatentIn 1.71
; SEQ ID NO 34
; LENGTH: 2510
; TYPE: DNA
; ORGANISM: Bacillus subtilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (461)..(2491)
US-10-433-577-34
Alignment Scores:
Pred. No.: 9,54e-16 Length: 2510
Score: 188.00 Matches: 34
Percent Similarity: 61.2% Conservative: 18
Best Local Similarity: 40.0% Mismatches: 31
Query Match: 40.1% Indels: 2
DB: 7 Gaps: 1

US-09-917-376-5 (1-88) x US-10-433-577-34 (1-2510)
QY 4 GlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnIleLysPro 23
Db 2051 GGTATTCTGTACATAACAGCAGCGGATGGAGTATGAACAGCAACCAATCCGTCG 2110
QY 24 GlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArg 43
Db 2111 CAGCTTCAATATAAATAACCGCAATACCGGTTGATTAAAGATGTCACTGCCCGT 2170
QY 44 TyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
Db 2171 TACTGGTATACGCGAAACAAAGGCCAAACGTT-----GACTGTGACTACGCGCAG 2224
QY 64 MetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
Db 2225 CTGGATCGCGCAATGTGACATACAAAGTTTGTGACGTTGCTAAACCAACCAAGAGTGCA 2284
QY 84 AspThrTyrLeuGln 88
Db 2285 GATACCTATCTGGAA 2299

```

```

RESULT 12
US-10-933-404-1
; Sequence 1, Application US/10933404
; Publication No. US20050118729A1
; GENERAL INFORMATION:
; APPLICANT: MORAG, Ely
; TITLE OF INVENTION: MICROARRAYS OF CELLULOSE BINDING CHIMERIC PROTEINS AND METHODS OF
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 85189-7100
; CURRENT APPLICATION NUMBER: US/10/933,404
; CURRENT FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: PCT/IL03/00177
; PRIOR FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US 60/362,061
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 501
; TYPE: DNA
; ORGANISM: Clostridium thermocellum
US-10-933-404-1

Alignment Scores:
Pred. No.: 2,13e-13 Length: 501
Score: 165.00 Matches: 35
Percent Similarity: 59.8% Conservative: 23
Best Local Similarity: 36.1% Mismatches: 29
Query Match: 35.2% Indels: 10
DB: 9 Gaps: 2

US-09-917-376-5 (1-88) x US-10-933-404-1 (1-501)

Qy 1 ValSerGlyValValGlnTyrLysAsnAspSerAlaProGlyAspAsnGln 20
Db 13 GTATCAGGCAATTTGAAGGTTGAATTTACACAGCAATCTTCAGATACCTAACTCA 72
Qy 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerThrLeuValTyrAsnCysAsp 60
Db 73 ATCAATCCTCAGTTCAGGTTACTAATACCGAGCAGCAAGATGCAATTTGTTCCAAACTC 132
Qy 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
Db 133 ACATTGAGATATTATTATATACAGTACGAGCAGCAAGATGCAATTTGTTCCAAACTC 189
Qy 61 TrpAlaAlaMet-----GlyCysGlyAsnIleArgAla 71
Db 190 CATGCTGCAATAATCGGCAGTACGAGCAGCTACAAACGGAATTTACTTCAAAATGTAAAGGA 249
Qy 72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
Db 250 ACATTGTTAAATGAGTTCTCTCAACAAATAACGAGACACCTTACTTGA 300

RESULT 13
US-10-372-054-3
; Sequence 3, Application US/10372054
; Publication No. US20030203466A1
; GENERAL INFORMATION:
; APPLICANT: Kauppinen, Markus
; APPLICANT: Schulein, Martin
; APPLICANT: Schnorr, Kirk
; APPLICANT: Andersen, Lene
; APPLICANT: Bjornvad, Mads
; TITLE OF INVENTION: Novel Mannanases
; FILE REFERENCE: 5440.204-US
; CURRENT APPLICATION NUMBER: US/10/372,054
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: US/09/339,159B
; PRIOR FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3

US-09-917-376-5 (1-88) x US-10-372-054-3 (1-1438)

Qy 1 ValSerGlyValValGlnTyrLysAsnAspSerAlaProGlyAspAsnGln 20
Db 940 GTATCAGGCAATTTGAAGGTTGAATTTACACAGCAATCTTCAGATACCTAACTCA 999
Qy 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 1000 ATCAATCCTCAGTTCAGGTTACTAATACCGAGCAGCAAGATGCAATTTGTTCCAAACTC 1059
Qy 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
Db 1060 ACATTGAGATATTATTATATACAGTACGAGCAGCAAGATGCAATTTGTTCCAAACTC 1116
Qy 61 TrpAlaAlaMet-----GlyCysGlyAsnIleArgAla 71
Db 1117 CATGCTGCAATAATCGGCAGTACGAGCAGCTACAAACGGAATTTACTTCAAAATGTAAAGGA 1176
Qy 72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
Db 1177 ACATTGTTAAATGAGTTCTCTCAACAAATAACGAGACACCTTACTTGA 1227

RESULT 14
US-10-072-152-11
; Sequence 11, Application US/10072152
; Publication No. US20020142438A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schulein, Martin
; APPLICANT: Lange, Niels E.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Moller, Soren
; APPLICANT: Glad, Sanne O. S.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Schnorr, Kirk
; APPLICANT: Kongsbak, Lars
; TITLE OF INVENTION: No. US20020142438A1el Pectate Lyases
; FILE REFERENCE: 5378.200-US
; CURRENT APPLICATION NUMBER: US/10/072,152
; CURRENT FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: US/09/198,955
; PRIOR FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: 1343/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 1344/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/067,249
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 60/067,240
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 09/073,684
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 09/184,217
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1482
; TYPE: DNA
; ORGANISM: Clostridium thermocellum

```


US-10-072-152-11

Alignment Scores:

Pred. No.: 9,04e-13 Length: 1482
Score: 165.00 Matches: 35
Percent Similarity: 59.8% Conservative: 23
Best Local Similarity: 36.1% Mismatches: 29
Query Match: 35.2% Indels: 10
DB: 5 Gaps: 2

US-09-917-376-5 (1-88) x US-10-072-152-11 (1-1482)

QY 1 ValSerGlyGlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspHngln 20
DB 991 GTATCAGGCAATTTGAAGGTTGAATTCACACAGCAATCCTTCAGATACTACTACTCA 1050
QY 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
DB 1051 ATCAATCCTCAGTTCAGGTTACTAATACCGAAGCAGTGCAATGATTGTCCAAACTC 1110
QY 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
DB 1111 ACATTGAGATATTATTATACAGTACGACGACAGAAAGATCAGACCTTCTGG--TGTGAC 1167
QY 61 TrpAlaAlaMet-----GlyCysGlyAsnIleArgala 71
DB 1168 CATGCTGCAATAATCGGCAGTAAACGCGACTACACGGAATTTACTTCAAAATGTAAAGGA 1227
QY 72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
DB 1228 ACATTGTGTAATAATGAGTTCTTCAACAAATAACGACGACACCTACTCTTGAA 1278

RESULT 15

US-10-655-433-11
; Sequence 11, Application US/10655433
; Publication No. US20040067572A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schulten, Martin
; APPLICANT: Lange, Niels E.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Moller, Soren
; APPLICANT: Glad, Samne O. S.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Schnort, Kirk
; APPLICANT: Kongsbak, Lars
; TITLE OF INVENTION: No. US20040067572A1el Pectate Lyases
; FILE REFERENCE: 5378.200-US
; CURRENT APPLICATION NUMBER: US/10/655,433
; CURRENT FILING DATE: 2003-09-04
; PRIOR APPLICATION NUMBER: US/09/198,955A
; PRIOR FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: 1343/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 1344/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/067,249
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 60/067,240
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 09/073,684
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 09/184,217
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID-NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1482
; TYPE: DNA
; ORGANISM: Clostridium thermocellum
US-10-655-433-11

Alignment Scores:

Pred. No.: 9,04e-13 Length: 1482
Score: 165.00 Matches: 35
Percent Similarity: 59.8% Conservative: 23
Best Local Similarity: 36.1% Mismatches: 29
Query Match: 35.2% Indels: 10
DB: 5 Gaps: 2

US-09-917-376-5 (1-88) x US-10-655-433-11 (1-1482)

QY 1 ValSerGlyGlyValLysValGlnTyrLysAsnAspSerAlaProGlyAspHngln 20
DB 991 GTATCAGGCAATTTGAAGGTTGAATTCACACAGCAATCCTTCAGATACTACTACTCA 1050
QY 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
DB 1051 ATCAATCCTCAGTTCAGGTTACTAATACCGAAGCAGTGCAATGATTGTCCAAACTC 1110
QY 41 ThrValArgTyrTrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAsp 60
DB 1111 ACATTGAGATATTATTATACAGTACGACGACAGAAAGATCAGACCTTCTGG--TGTGAC 1167
QY 61 TrpAlaAlaMet-----GlyCysGlyAsnIleArgala 71
DB 1168 CATGCTGCAATAATCGGCAGTAAACGCGACTACACGGAATTTACTTCAAAATGTAAAGGA 1227
QY 72 SerPheGlySerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
DB 1228 ACATTGTGTAATAATGAGTTCTTCAACAAATAACGACGACACCTACTCTTGAA 1278

Search completed: March 2, 2006, 20:26:36
Job time : 701.153 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 2, 2006, 12:24:59 ; Search time 94.4333 Seconds
(without alignments)
2043.020 Million cell updates/sec

Title: US-09-917-376-5

Perfect score: 469

Sequence: 1 VSGGVKVOYKNDSPAGDQ.....IRASFGSVNPTPTADTYLQ 88

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7218512 seqs, 1096188070 residues

Total number of hits satisfying chosen parameters: 14437024

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame p2n.model -DEV=xlp
-Q=/abs/ABSWEB/spool/US0917376/runat_02032006_091502_8293/app_query.fasta.1
-DB=Published Applications NA New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abs03p
-USER=US0917376.ecgn_1_1067 @runat_02032006_091502_8293 -NCPU=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New.*

1: /cgn2_6/ptodata/2/pubnpa/US08_NEW_PUB.seq.*
2: /cgn2_6/ptodata/2/pubnpa/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubnpa/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubnpa/PCT_NEW_PUB.seq.*
5: /cgn2_6/ptodata/2/pubnpa/US05_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubnpa/US09_NEW_PUB.seq.*
7: /cgn2_6/ptodata/2/pubnpa/US10_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubnpa/US10_NEW_PUB.seq.*
9: /cgn2_6/ptodata/2/pubnpa/US11_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubnpa/US11_NEW_PUB.seq.*
11: /cgn2_6/ptodata/2/pubnpa/US11_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubnpa/US11_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubnpa/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	163	34.8	1500	8	US-10-517-939-327
2	157	33.5	2602	8	US-10-510-386-229
3	113	24.1	1146	12	US-11-179-411-21
4	113	24.1	1146	12	US-11-175-766-21

5	108	23.0	2217	8	US-10-517-939-343	Sequence 343, App
6	70	14.9	15876	12	US-11-052-554A-660	Sequence 660, App
7	69.5	14.8	617	6	US-09-925-065A-353393	Sequence 353393, A
8	68.5	14.6	1084	8	US-10-750-185-31182	Sequence 31182, A
9	68.5	14.6	1084	8	US-10-750-623-31182	Sequence 31182, A
10	68.5	14.6	2231	9	US-11-072-512-855	Sequence 855, App
11	68.5	14.6	4952	12	US-11-136-527-487	Sequence 487, App
12	68	14.5	2268	8	US-10-517-939-329	Sequence 329, App
13	68	14.5	35344	8	US-10-995-561-13307	Sequence 13307, A
14	67.5	14.4	556	6	US-09-925-065A-228751	Sequence 228751, A
15	67.5	14.4	2564	9	US-11-072-512-279	Sequence 279, App
16	67	14.3	523	6	US-09-925-065A-39541	Sequence 39541, A
17	67	14.3	17186	12	US-11-121-086-105	Sequence 105, App
18	66.5	14.2	600	8	US-10-750-185-1174	Sequence 1174, App
19	66.5	14.2	600	8	US-10-750-623-1174	Sequence 1174, App
20	66.5	14.2	1847	12	US-11-000-688-1099	Sequence 1099, App
21	66.5	14.2	2565	8	US-10-750-185-48986	Sequence 48986, A
22	66.5	14.2	2565	8	US-10-750-623-48986	Sequence 48986, A
23	65.5	14.0	617	6	US-09-925-065A-353394	Sequence 353394, A
24	65.5	14.0	1218	6	US-09-925-065A-547306	Sequence 547306, A
25	65.5	14.0	1218	6	US-09-925-065A-547307	Sequence 547307, A
26	65.5	14.0	1218	6	US-09-925-065A-547308	Sequence 547308, A
27	64.5	13.8	600	6	US-09-925-065A-336480	Sequence 336480, A
28	64.5	13.8	600	6	US-09-925-065A-336481	Sequence 336481, A
29	64.5	13.8	601	6	US-09-925-065A-729386	Sequence 729386, A
30	64.5	13.8	2370	6	US-09-925-065A-91998	Sequence 91998, A
31	64.5	13.8	2370	6	US-09-925-065A-91999	Sequence 91999, A
32	64.5	13.8	159138	8	US-10-995-561-13230	Sequence 13230, A
33	64	13.6	556	6	US-09-925-065A-389647	Sequence 389647, A
34	64	13.6	585	6	US-09-925-065A-385747	Sequence 385747, A
35	64	13.6	611	6	US-09-925-065A-658978	Sequence 658978, A
36	64	13.6	659	6	US-09-925-065A-87693	Sequence 87693, A
37	64	13.6	659	6	US-09-925-065A-87694	Sequence 87694, A
38	64	13.6	1898	9	US-11-072-512-1678	Sequence 1678, App
39	64	13.6	1990	12	US-11-062-225-2	Sequence 2, Appli
40	64	13.6	1990	12	US-11-062-225-10	Sequence 10, Appl
41	64	13.6	3074	12	US-11-000-688-315	Sequence 315, App
42	63.5	13.5	596	6	US-09-925-065A-521812	Sequence 521812, A
43	63.5	13.5	597	6	US-09-925-065A-258810	Sequence 258810, A
44	63.5	13.5	597	6	US-09-925-065A-258811	Sequence 258811, A
45	63.5	13.5	606	6	US-09-925-065A-484851	Sequence 484851, A

ALIGNMENTS

RESULT 1

US-10-517-939-327
; Sequence 327, Application US/10517939
; Publication No. US20060003433A1
; GENERAL INFORMATION:
; APPLICANT: Steer, Brian
; APPLICANT: Callen, Walter
; APPLICANT: Healey, Shaun
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Wu, Di
; APPLICANT: Blum, David
; APPLICANT: Esteghalian, Alireza
; TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM
; FILE REFERENCE: 564462007901
; CURRENT APPLICATION NUMBER: US/10/517,939
; CURRENT FILING DATE: 2004-12-13
; PRIOR APPLICATION NUMBER: PCT/US03/19153
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/389,299
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 380
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 327
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:

OTHER INFORMATION: Obtained from an environmental sample.
US-10-517-939-327

Alignment Scores:

Pred. No.: 8.87e-11 Length: 1500
Score: 163.00 Matches: 30
Percent Similarity: 60.0% Conservatives: 21
Best Local Similarity: 35.3% Mismatches: 30
Query Match: 34.8% Indels: 4
DB: 8 Gaps: 2

US-09-917-376-5 (1-88) x US-10-517-939-327 (1-1500)

Qy 5 VallysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnLeuLysProGly 24
Db 1060 CTATCTGTGCAATACAGACAGGGGATGTAACACCAACCAATCCGCTCAG 1119
Qy 25 LeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValValArgTyr 44
Db 1120 ATCATGTGAAACAAACAGCAAGCACCCTTAATTTAAATAATGTAACCTGCCGTAC 1179
Qy 45 TrpPhe---ThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
Db 1180 TGGTATAACAGAAACAAAGGCCAAAC-----TTGACTGTGACTAGCGGAAG 1230
Qy 64 MetGlyCysGlyAsnLeuArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
Db 1231 ATCGGATCGACCAATGTGACGACCAAGTTGTGACATTACAAACAACTGTAAGAGGTGCA 1290
Qy 84 AspThrTyrLeuGln 88
Db 1291 GATGCTATCTGGAA 1305

RESULT 2

US-10-510-386-229
; Sequence 229, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510.386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 229
; LENGTH: 2602
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (501)..(2102)
US-10-510-386-229

Alignment Scores:

Pred. No.: 9.84e-10 Length: 2602
Score: 157.00 Matches: 32
Percent Similarity: 56.5% Conservatives: 16
Best Local Similarity: 37.6% Mismatches: 33
Query Match: 33.5% Indels: 4
DB: 8 Gaps: 2

US-09-917-376-5 (1-88) x US-10-510-386-229 (1-2602)

Qy 5 VallysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnLeuLysProGly 24
Db 1665 ATAGCAGTACATACAGAGGGGGACCAACATGTAAACGGCAACCAATCCGCTCAG 1724
Qy 25 LeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValValArgTyr 44

Db 1725 CTCACATTAAAAACACAGCAAAAAACCGGTCTTTAAATCGAATCACTGTCCGCTAC 1784
Qy 45 TrpPhe---ThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAla 63
Db 1785 TGGTATAAACGAATCGCAAGGACAAAT-----TTGACTGTGACTATGCCCAA 1835
Qy 64 MetGlyCysGlyAsnLeuArgAlaSerPheGlySerValAsnProAlaThrProThrAla 83
Db 1836 ATCGGCTGCAGCAAAATCAGCACAATTCGTCAATTAAAAAAGCGGTAAACGGAGCA 1895
Qy 84 AspThrTyrLeuGln 88
Db 1896 GACACGTATCTGAA 1910
RESULT 3
; Sequence 21, Application US/11179411
; Publication No. US20050266491A1
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: Szczepaniak, William
; TITLE OF INVENTION: RENILLA RENIFORMIS FLOURESCENT PROTEINS, NUCLEIC ACIDS ENCODING TH
; TITLE OF INVENTION: FLOURESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIGH TH
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS
; FILE REFERENCE: 24729-128
; CURRENT APPLICATION NUMBER: US/11/179,411
; CURRENT FILING DATE: 2005-07-12
; PRIOR APPLICATION NUMBER: US/09/808,898
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/189,691
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 09/277,716
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 08/757,046
; PRIOR FILING DATE: 1996-11-25
; PRIOR APPLICATION NUMBER: 08/597,274
; PRIOR FILING DATE: 1996-02-06
; PRIOR APPLICATION NUMBER: 08/908,909
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 08/990,103
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 1146
; TYPE: DNA
; ORGANISM: Gaussia
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1146)
; OTHER INFORMATION: Nucleotide sequence encoding a CBD-Gaussia
; OTHER INFORMATION: luciferase fusion protein
US-11-179-411-21
Alignment Scores:
Pred. No.: 0.00017 Length: 1146
Score: 113.00 Matches: 29
Percent Similarity: 53.2% Conservatives: 21
Best Local Similarity: 30.9% Mismatches: 32
Query Match: 24.1% Indels: 12
DB: 12 Gaps: 4
US-09-917-376-5 (1-88) x US-11-179-411-21 (1-1146)
Qy 5 VallysValGlnTyrLysAsnAspSerAlaProGlyAspAsnGlnLeuLysProGly 24
Db 1 ATGTCAGTTGAATTTTCACTTAACAAATACGACACAAACAACTCAATTAACCAATA 60
Qy 25 LeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValValArgTyr 44
Db 61 ATCAAAATTACTAACAACATCTGACAGTCAATTAAATTTAAATGACGTAAAGCTTAGATAT 120

QY 45 TrpPheThrArgAspGlyGlySerThrLeuValTyrAsnCysAspTrpAlaMet 64
 Db 121 TATTACAAAGTGGTACACAGCAAACTTTCTGG---TGTGACCATGCTGGTGA 177
 QY 65 GlyCysGlyAsn-----IleArgAlaSerPhe-----Gly 74
 Db 178 TTATTAGGAATAGTATGTTGATACACTAGCAAGTACAGCAAACTTCGTTAAAGNA 237
 QY 75 SerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
 Db 238 ACAGCAAGCCCA---ACATCAACCTATGATACATATGTTGAA 276

RESULT 4

US-11-175-766-21
 ; Sequence 21, Application US/11175766
 ; Publication No. US2005027211A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bryan, Bruce
 ; APPLICANT: Szent-Gyorgyi, Christopher
 ; APPLICANT: Szczepaniak, William
 ; TITLE OF INVENTION: RENILLA RENIFORMIS FLOURESCENT PROTEINS, NUCLEIC ACIDS ENCODING T
 ; TITLE OF INVENTION: FLOURESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIGH TH
 ; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS
 ; FILE REFERENCE: 24729-128
 ; CURRENT APPLICATION NUMBER: US/11/175,766
 ; CURRENT FILING DATE: 2005-07-06
 ; PRIOR APPLICATION NUMBER: US/09/808,898
 ; PRIOR FILING DATE: 2001-03-15
 ; PRIOR APPLICATION NUMBER: 60/189,691
 ; PRIOR FILING DATE: 2000-03-15
 ; PRIOR APPLICATION NUMBER: 09/277,716
 ; PRIOR FILING DATE: 1999-03-26
 ; PRIOR APPLICATION NUMBER: 08/757,046
 ; PRIOR FILING DATE: 1996-11-25
 ; PRIOR APPLICATION NUMBER: 08/597,274
 ; PRIOR FILING DATE: 1996-02-06
 ; PRIOR APPLICATION NUMBER: 08/908,909
 ; PRIOR FILING DATE: 1997-08-08
 ; PRIOR APPLICATION NUMBER: 08/990,103
 ; PRIOR FILING DATE: 1997-12-12
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 21
 ; LENGTH: 1146
 ; TYPE: DNA
 ; ORGANISM: Gaussia
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(1146)
 ; OTHER INFORMATION: Nucleotide sequence encoding a CBD-Gaussia
 ; OTHER INFORMATION: luciferase fusion protein
 US-11-175-766-21

Alignment Scores:

Pred. No.:	0.00017	Length:	1146
Score:	113.00	Matches:	29
Percent Similarity:	53.2%	Conservative:	21
Best Local Similarity:	30.9%	Mismatches:	32
Query Match:	24.1%	Indels:	12
DB:	12	Gaps:	4

US-09-917-376-5 (1-88) x US-11-175-766-21 (1-1146)

QY 5 ValGlnTyrLysAsnAsnSerAlaProGlyAspAsnGlnIleLysProGly 24
 Db 1 ATGTCAGTTGAATTTACAACTCTAACAAATCAGCAAAACAACTCAATTACACCAATA 60
 QY 25 LeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArgTyr 44
 Db 61 ATCAAAATTACTAACACATCTGACAGTGAATTTAAATTTAAATGACGTAAAGTTAGATAT 120
 QY 45 TrpPheThrArgAspGlyGlySerThrLeuValTyrAsnCysAspTrpAlaMet 64

Db 121 TATTACAAAGTGGTACACAGCAAACTTTCTGG---TGTGACCATGCTGGTGA 177
 QY 65 GlyCysGlyAsn-----IleArgAlaSerPhe-----Gly 74
 Db 178 TTATTAGGAATAGTATGTTGATACACTAGCAAGTACAGCAAACTTCGTTAAAGNA 237
 QY 75 SerValAsnProAlaThrProThrAlaAspThrTyrLeuGln 88
 Db 238 ACAGCAAGCCCA---ACATCAACCTATGATACATATGTTGAA 276

RESULT 5

US-10-517-939-343
 ; Sequence 343, Application US/10517939
 ; Publication No. US2006000343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Steer, Brian
 ; APPLICANT: Callen, Walter
 ; APPLICANT: Healey, Shaun
 ; APPLICANT: Hazlewood, Geoff
 ; APPLICANT: Wu, Di
 ; APPLICANT: Blum, David
 ; APPLICANT: Esteghlalian, Alireza
 ; TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM
 ; TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
 ; FILE REFERENCE: 564462007901
 ; CURRENT APPLICATION NUMBER: US/10/517,939
 ; CURRENT FILING DATE: 2004-12-13
 ; PRIOR APPLICATION NUMBER: PCT/US03/19153
 ; PRIOR FILING DATE: 2003-06-16
 ; PRIOR APPLICATION NUMBER: 60/389,299
 ; PRIOR FILING DATE: 2002-06-14
 ; NUMBER OF SEQ ID NOS: 380
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 343
 ; LENGTH: 2217
 ; TYPE: DNA
 ; ORGANISM: Unknown
 ; FEATURE:
 ; OTHER INFORMATION: Obtained from an environmental sample.
 US-10-517-939-343

Alignment Scores:

Pred. No.:	0.00159	Length:	2217
Score:	108.00	Matches:	27
Percent Similarity:	53.6%	Conservative:	18
Best Local Similarity:	32.1%	Mismatches:	31
Query Match:	23.0%	Indels:	8
DB:	8	Gaps:	4

US-09-917-376-5 (1-88) x US-10-517-939-343 (1-2217)

QY 7 ValGlnTyrLysAsnAsnSerAlaProGlyAspAsn-----GlnIleLysProGly 24
 Db 1771 CTTCACTACCGC-----TCCGCTGATGGAGATAACAACCTATCAAAATGAAGCCTCAG 1821
 QY 25 LeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArgTyr 44
 Db 1822 TTCACGATCAAGAACGACGAGCAAGTCCCATCCCGTTAAGCGAGTGACGATCCGCTAC 1881
 QY 45 TrpPheThrArgAspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaMet 64
 Db 1882 TATTCACCGCGGAG---AGCAGCGAGCGGTGGATACAGGATCGACTGGGCCCAATTC 1938
 QY 65 GlyCysGlyAsnIleArgAlaSerValAsnProAlaThrProThrAlaMet 84
 Db 1939 GGAGCAGAGCATGTCACAGACGACGCTC-----GTTCCGCCATCCGATCGCGGCGCAC 1992
 QY 85 ThrTyrLeuGln 88
 Db 1993 GCCTATGTCGAG 2004

RESULT 6

```
US-11-052-554A-660
; Sequence 660, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 660
; LENGTH: 15876
; TYPE: DNA
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-660

Alignment Scores:
Pred. No.: 1,15e+03 Length: 15876
Score: 70.00 Matches: 29
Percent Similarity: 50.6% Conservative: 15
Best Local Similarity: 33.3% Mismatches: 19
Query Match: 14.9% Indels: 24
DB: 12 Gaps: 7

US-09-917-376-5 (1-88) x US-11-052-554A-660 (1-15876)
QY 17 GlyAsp-----AsnGlnIleLysProGlyLeuGlnLeuValAsnThrGlySerSer 33
Db 2542 GGTGATGATGATCAAGCGCGTCGAAAGAGGCGGCGTGGTAAAGCGGAGCACC 2601
QY 34 Ser---ValAspLeu---SerThrValThrValArgTyrTrpPheThrArgAspGlyGly 51
Db 2602 AGCGGTGTCGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2646
QY 52 -----SerSerThrLeuValThrAsnCysAspTrp----- 61
Db 2647 AAAAATATACACACACAGTGAAGCTAAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 2706
QY 62 -----AlaAlaMet-----GlyCysGlyAsnIleArgAlaSerPheGlySerValAsn 77
Db 2707 GATCTGCTGCTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2766
QY 78 ProAlaThrProThrAlaAsp 84
Db 2767 GGCAACAGCGCGCCAGCGCGAT 2787

RESULT 7
US-09-925-065A-353393
; Sequence 353393, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09

; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 353393
; LENGTH: 617
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-353393

Alignment Scores:
Pred. No.: 31.8 Length: 617
Score: 69.50 Matches: 19
Percent Similarity: 44.8% Conservative: 11
Best Local Similarity: 28.4% Mismatches: 20
Query Match: 14.8% Indels: 17
DB: 6 Gaps: 2

US-09-917-376-5 (1-88) x US-09-925-065A-353393 (1-617)
QY 25 LeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrValArgTyr 44
Db 214 TTGAGAGCAGTTAAGGAGGGTCAAGATCTTTAGGCTCAGCTCCATGACTCTAGGCCA 273
QY 45 TrpPheThrArgAspGlyGlySerSerThrLeuValThrAsnCysAspTrpAlaAla--- 63
Db 274 TGGTTT-----CTAATCTTATGGCTAGTTTCTTGGTCTGCTGCC 312
QY 64 -----MetGlyCysGlyAsnIleArgAlaSerPheGly 74
Db 313 CAGGCAAGAGGTAAAGCTATCTTAAGAGGGGCTGTACCACTTGTGTTTACTACTATAA 372
QY 75 SerValAsnProAlaThrPro 81
Db 373 ACTATAAACCAGCTCCTCCA 393

RESULT 8
US-10-750-185-31182
; Sequence 31182, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 31182
; LENGTH: 1084
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-31182

Alignment Scores:
Pred. No.: 81.8 Length: 1084
Score: 68.50 Matches: 25
Percent Similarity: 40.2% Conservative: 12
Best Local Similarity: 27.2% Mismatches: 32
Query Match: 14.6% Indels: 23
DB: 8 Gaps: 5

US-09-917-376-5 (1-88) x US-10-750-185-31182 (1-1084)
QY 16 ProGlyAspAsnGlnIleLysProGlyLeuGlnLeuVal-----AsnThr 30
Db 738 CCAGGACCAC-----CCCAATTCACCTGGTGTCTCTCTCTCTCTCTCTCTCTCTCT 788
```

Qy	31	GlySerSerValAlaSerThrValThrValArgTyrTrpPheThrArgAspGly	50
Db	789	GGGCATTTAGACATCCAGAAGCTCAGTCCTCGCCATCAGACATTTGGCTAAAGTGCACATGGG	848
Qy	51	GlySerSerThrLeuValTyrAsnCysAspTrpAlaAlaMetGlyCysGlyAsnIleArg	70
Db	849	GGTGCAGAAG--ATGACACTGACGTGCGCCAGCGAGAGCTGGGAGGGCTGGTTTCAGG	905
Qy	71	Ala-----SerPheGlySerValAsnPro---	78
Db	906	ACCCACGCCTTCACTCCACAGTTTACACAGCGAGCTCTGGCGCGTGATCTCTCAG	965
Qy	79	-----AlaThrProThrAlaAspThrTyrLeuGln	88
Db	966	GAAGTGGCCATCCCGAGGAGGAGCTGTTGGCTCGAG	1001

```

RESULT 10
US-11-072-512-855
; Sequence 855, Application US/11072512
; Publication No. US2006029945A1
; GENERAL INFORMATION: TAKAO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOFUKU
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,000
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-370,000
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 855
; LENGTH: 2231
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-072-512-855

```

```

; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 487
; LENGTH: 4952
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-487

Alignment Scores:
Pred. No.: 470 Length: 4952
Score: 68.50 Matches: 22
Percent Similarity: 44.6% Conservative: 7
Best Local Similarity: 33.8% Mismatches: 25
Query Match: 14.6% Indels: 11
DB: 12 Gaps: 3

US-09-917-376-5 (1-88) x US-11-136-527-487 (1-4952)
QY 31 GlySerSerValAspLeuSerThrValThrValArgTyrTrpPheThrArgAspGly 50
Db 1413 GGGACCCAGAACAGAGCTAGAGTCTAGTGCAGCACTTATACACTCAAGGCAGATGGA 1472
QY 51 -----GlySerSer-----ThrLeuValTyrAsnCysAsp----- 60
Db 1473 ACCCAGTCTCCAGTGGTCTTCTGGAACACAGCATCTACAGTGTGAGTCTGCACG 1532
QY 61 ---TrpAlaAlaMetGlyCysGlyAsnIleArgAlaSerPheGlySerValAsnProAla 79
Db 1533 GTCTAGGAGGCAAGGAGTGAAGAACATAGCCGTGACCTTACCTCTGTAGCCAACTTA 1592
QY 80 ThrProThrAlaAsp 84
Db 1593 ACAATAACCCCGGAC 1607

RESULT 12
US-10-517-939-329
; Sequence 329, Application US/10517939
; Publication No. US20060003433A1
; GENERAL INFORMATION:
; APPLICANT: Steer, Brian
; APPLICANT: Callen, Walter
; APPLICANT: Healey, Shaun
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Wu, Di
; APPLICANT: Blum, David
; APPLICANT: Esteghlalian, Alireza
; TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM
; FILE REFERENCE: 564462007901
; CURRENT APPLICATION NUMBER: US/10/517,939
; PRIOR FILING DATE: 2004-12-13
; PRIOR APPLICATION NUMBER: PCT/US03/19153
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/389,299
; NUMBER OF SEQ ID NOS: 380
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 329
; LENGTH: 2268
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample.
US-10-517-939-329

Alignment Scores:
Pred. No.: 222 Length: 2268
Score: 68.00 Matches: 31
Percent Similarity: 37.9% Conservative: 8
Best Local Similarity: 30.1% Mismatches: 40
Query Match: 14.5% Indels: 24
DB: 8 Gaps: 5

US-09-917-376-5 (1-88) x US-10-517-939-329 (1-2268)
QY 2 SerGlyGlyVallysValGlnTyrLys-----AsnAsnAspSerAlaProGlyAspAsnGln 20
Db 1951 TCAGGGCGCTGCACGCTGCCTACGCCATCACCACGATCGGGCAGCGT----- 2001
QY 21 IleLysProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrVal 40
Db 2002 TTCACCGCCCAACGTTTACCTTCCCGCAATCAACCATCAGCAACGCTGGACGGAACGGCC 2055
QY 41 ThrValArgTyrTrpPheThrArgAsp----- 49
Db 2056 ACCCTGGCCTATGCTTTCCTCCCGCAATCAACCATCAGCAACGCTGGACGGAACGGCC 2115
QY 50 -----GlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAlaMetGlyCysGly 67
Db 2116 GTTCAGTCCGGCAGCAGCGTCAGCGTCACCAACGCCGCTTGAATGGCAGCCTGCCGCC 2175
QY 68 AsnIleArgAlaSerPheGly-----SerValAsnProAlaThrProThr 82
Db 2176 AACGTCCTCCGCGAGCTTTGGCTTCCAGGCGGAGCTACAGCGGCAATAACAGCGTCCCTGCC 2235
QY 83 AlaAspThr 85
Db 2236 AGCTTTACG 2244

RESULT 13
US-10-995-561-13307/c
; Sequence 13307, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13307
; LENGTH: 35344
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13307

Alignment Scores:
Pred. No.: 5,23e+03 Length: 35344
Score: 68.00 Matches: 16
Percent Similarity: 57.1% Conservative: 4
Best Local Similarity: 45.7% Mismatches: 11
Query Match: 14.5% Indels: 4
DB: 8 Gaps: 1

US-09-917-376-5 (1-88) x US-10-995-561-13307 (1-35344)
QY 51 GlySerSerThrLeuValTyr-----AsnCysAspTrpAlaAlaMetGlyCys 66
Db 23968 GGTAAAGACCACTTCCCTTATTCACCTTGAAGTCACTGTGAGTGGCGGCTTTTGGC 23909
QY 67 GlyAsnIleArgAlaSerPheGlySerValAsnProAlaThrPro 81
Db 23908 TTTAGAGTAGGGCATCATTTACCACCTTCTTTCCCTCTAGTCTCT 23864

RESULT 14
US-09-925-065A-228751/c
; Sequence 228751, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome

```

```
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 228751
LENGTH: 556
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-228751

Alignment Scores:
Pred. No.: 51          Length: 556
Score: 67.50          Matches: 25
Percent Similarity: 37.6%      Conservative: 7
Best Local Similarity: 29.4%    Mismatches: 20
Query Match: 14.4%             Indels: 33
DB: 6                       Gaps: 4

US-09-917-376-5 (1-88) x US-09-925-065A-228751 (1-556)
QY 23 ProGlyLeuGlnLeuValAsnThrGlySer----- 32
Db 238 CCTGGACTCCACCTTGTGCATACAGCAGTTAAGGAATGCCCTTCTGCACAAAGACTTTT 179
QY 33 -----SerSerValAspLeuSerThrValThrValArgTyrTrpPheThrArg 48
Db 178 ATTTTAGCTTAGTTTCTCTGGAATTTGTTTGGAGATTGAGTTACTGTGTAATATGG 119
QY 49 AspGlyGlySerSerThrLeuValTyrAsnCysAspTrpAlaAlaMet----- 64
Db 118 GACCCGGGG-----TAGGACTGGCTGGCGGCTCGCATCCTTCC 80
QY 65 -----GlyCysGlyAsnLeuArg---AlaSerPheGlySerVal 76
Db 79 TCCGAGGGCCCTGGGGAGAGGGTGTGGAGTGGCCCGGTGGCTGTTTCTCCCCCAT 20
QY 77 AsnProAlaThrPro 81
Db 19 TCCTGTTGGACCCCT 5

RESULT 15
US-11-072-279/c
Sequence 279, Application US/11072512
Publication No. US20060029945A1
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHICO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOKYUKI
APPLICANT: NAGAHARI, KENJI
```

```
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: Novel full length cDNA
FILE REFERENCE: 084335-0191
CURRENT APPLICATION NUMBER: US/11/072,512
CURRENT FILING DATE: 2005-03-07
PRIOR APPLICATION NUMBER: US 60/350,978
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: JP 2001-379298
PRIOR FILING DATE: 2001-11-05
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 279
LENGTH: 2564
TYPE: DNA
ORGANISM: Homo sapiens
US-11-072-512-279

Alignment Scores:
Pred. No.: 296          Length: 2564
Score: 67.50          Matches: 23
Percent Similarity: 42.1%      Conservative: 9
Best Local Similarity: 30.3%    Mismatches: 27
Query Match: 14.4%             Indels: 17
DB: 3                       Gaps: 3

US-09-917-376-5 (1-88) x US-11-072-512-279 (1-2564)
QY 23 ProGlyLeuGlnLeuValAsnThrGlySerSerValAspLeuSerThrValThrVal 42
Db 454 CCTCCGAAGCGATGGCTGACACTCGCTCCAGCAGCTTGTCCGCGAGCTCATCGAATACC 395
QY 43 ArgTyrTrpPheThr----- 47
Db 394 GACTGGTGAATCCAGCCGCGGCTCCCGTCGAGGTCCAGGAAGGGGGGCTGACTGA 335
QY 48 ArgAspGlyGlySerSerThrLeu---ValTyrAsnCysAspTrpAlaAlaMetGlyCys 66
Db 334 AGAGATGGCAGCAGCACACCATTTCTCTGTCTGGAACCTGCTCGATGGCTTGAGCGGCTCC 275
QY 67 GlyAsn---IleArgAlaSerPheGlySerValAsnProAlaThrPro 81
Db 274 GTGCAGTTGGTCAGGGTCTCTTTCAGGTCTCTCGCCCGTGGCCACGCC 227

Search completed: March 2, 2006, 13:23:29
Job time : 101.433 secs
```